

UNIVERSITE CATHOLIQUE DE LOUVAIN
Faculté de psychologie et des sciences de l'éducation.

Dominique Morsomme

**CONTRIBUTION A LA DETERMINATION
DE PARAMETRES SUBJECTIFS ET
OBJECTIFS POUR L'ETUDE DE LA
VOIX.**

Promoteur: Professeur M. Remacle

Co-Promoteurs : Professeur J. Costermans

Professeur M-A. Schelstraete

Louvain- la- Neuve 2001

"Où, le cœur n'est pas, il ne saurait y avoir de musique."

Piotr Ilitch Tchaïkovsky (1840-1893).

Remerciements

Au Professeur M. Crochet, Recteur Magnifique de l'Université Catholique de Louvain, au Professeur R. Lauwerys, Prorecteur, au Professeur M. Hupet, Doyen de la Faculté de Psychologie et des Sciences de l'Éducation, au Professeur D. Moulin, Doyen de la Faculté de Médecine, au Professeur M. Gersdorff, Chef du Département Clinique N.O.P.S., j'exprime ma reconnaissance pour m'avoir donné l'occasion de mener à bien mes travaux au sein de l'Université Catholique de Louvain.

Mon respect et ma profonde gratitude vont particulièrement au Professeur M. Remacle qui m'a accordé sa confiance en m'inculquant rigueur et précision ; je lui dois les connaissances théoriques et techniques nécessaires à l'élaboration de mes travaux. Qu'il soit remercié de m'avoir laissée libre de toutes initiatives dans le domaine du chant, ce qui m'a permis d'unir passion et vie professionnelle,

Ma gratitude va aussi au Professeur J. Costermans, inopinément décédé, qui a accepté d'être co-promoteur de ma thèse et m'a invitée à poursuivre plus spécifiquement mes travaux sur la construction d'une échelle perceptuelle auquel il s'intéressait vivement non seulement en sa qualité de scientifique mais également de violoniste averti,

J'exprime également ma reconnaissance au Professeur M-A. Schelstraete pour ses encouragements, ses conseils et sa disponibilité, ainsi qu'au Professeur M-P. de Partz de Courtray, pour ses commentaires judicieux et le temps consacré à la lecture du manuscrit ; au Professeur M. De Bodt et au Professeur F. Wuyts qui m'ont fait bénéficier sans restriction de leurs compétences et de leur savoir.

Mes remerciements vont encore,

au Docteur J. Jamart, Responsable du Centre de Biostatistique et de Documentation Médicale aux Cliniques Universitaires de Mont-godinne, qui a assuré le suivi statistique de mes travaux ;

à Monsieur W. Pirie, Artiste Lyrique, pour les nombreux échanges qui n'ont eu de cesse d'enrichir ma pratique logopédique,

au Professeur F. Estienne, qui m'a communiqué sa passion de la pratique logopédique en m'ouvrant les portes de son savoir. Elle a également accepté de relire avec minutie ce travail. Je lui sais gré de son dévouement et de ses perpétuels encouragements,

à Madame M. Duchêne, Philologue romaniste qui m'a prodigué de judicieux conseils lors la rédaction de ce travail,

à Monsieur C. Deneffe, photographe, pour son dévouement et son efficacité dans la mise en page et l'illustration ;

aux Professeurs de chant et à leurs élèves, aux étudiants des écoles de la Communauté Française, à l'ensemble des personnes dont la participation active aux études m'a permis d'évoluer dans mes recherches et particulièrement à Marie-Cécile Jamar, Evelyne Artus, Andrée Orban, Catherine de Menten et Ombeline d'Oultremont,

à mes collègues du Centre d'audiophonologie Saint-Luc pour leur soutien moral dans les moments difficiles.

A mon mari, pour sa patience, ses encouragements et son amour infiniment grand,

à ma famille, et particulièrement à ma sœur pour avoir partagé mon enthousiasme et mes déceptions dans la réalisation de ce projet,

à Joelle qui m'a permis de relativiser les événements de la vie et d'en apprécier chaque instant,

que chacun d'eux trouvent ici le témoignage de ma profonde reconnaissance.

TABLE DES MATIÈRES

INTRODUCTION

CHAPITRE I : L'ANATOMOPHYSIOLOGIE

1.1. L'anatomie	p17
1.1.1. Les organes de la phonation	p17
1.1.2. Les ligaments	p18
1.1.3. Les muscles intrinsèques	p19
1.1.4. Les muscles extrinsèques	p21
1.1.5. La structure des cordes vocales	p22
1.1.6. Le tractus vocal supra-glottique	p24
1.1.7. Le tractus vocal sous-glottique	p24
1.2. La physiologie	p25
1.2.1. Le Mécanisme vibratoire des cordes vocales	p25
1.2.2. Les Caractéristiques acoustiques de la voix	p28
1.2.2.1. <u>Définitions</u>	p28
1.2.2.2. <u>La hauteur</u>	p29
1.2.2.2.1 <i>Les Facteurs qui affectent la hauteur</i>	p32
1.2.2.2.1.1. L'âge	p32
1.2.2.2.1.2. La pathologie	p40
1.2.2.2.1.3. La culture	p42
1.2.2.2.1.4. L'état émotionnel	p43
1.2.2.2.1.5. Les facteurs hormonaux	p45
1.2.2.2.1.6. Le sexe	p47
1.2.2.2.1.7. La posture	p59
1.2.2.3. <u>L'Intensité vocale</u>	p51
1.2.2.3.1. <i>Le Mécanisme de contrôle</i>	p51
1.2.2.3.2. <i>La perception de l'intensité</i>	p51

1.2.2.3.3. <i>L'intensité en fonction de la fréquence</i>	p53
1.2.2.3.4. <i>L'effet Lombard</i>	p53
1.2.2.3.5. <i>Le niveau de l'intensité</i>	p55
1.2.2.4. <u>Le timbre</u>	p56
1.2.2.5. <u>L'Application à la voix chantée</u>	p61
1.2.2.6. <u>La vibration glottique</u>	p64
1.2.2.6.1. <i>L'irrégularité vibratoire</i>	p64
1.2.2.6.2. <i>Les mesures de la perturbation : Le Jitter, le Shimmer, le Rapport signal/bruit</i>	p66
1.2.2.6.3. <i>La validité des mesures</i>	p67
1.2.2.6.4. <i>Les corrélations de mesures de perturbation avec la perception</i>	p71

CHAPITRE II : LA VIDEOSTROBOSCOPIE

2.1. La vidéostroboscopie	p75
2.1.1. La description de la technique	p75
2.1.2. Les limites de la vidéostroboscopie	p82
2.1.3. L'examen du sujet	p82
2.1.4. La fiabilité du diagnostic	p83

CHAPITRE III : L'EVALUATION SUBJECTIVE

3.1. Les échelles d'auto-évaluation	p87
3.2.1. L'échelle bipolaire d'auto-estimation vocale de Dejonckere	p87
3.2.2. L'index de handicap vocal ou V.H.I. (Voice Handicap Index) de Jacobson	p88
3.2. Les échelles perceptuelles	p90
3.2.1. Les profils décrits dans la littérature	p90
3.2.1.1. <u>Le "Buffalo Voice Profile" de Wilson</u>	p91
3.2.1.2. <u>Le schéma de Hammarberg</u>	p92

3.2.1.3. <u>Le profil vocal de Laver</u>	p94
3.2.1.4. <u>L'échelle G.R.B.A.S. de Hirano</u>	p96
3.2.2 La fiabilité de l'échelle	p98
3.2.2.1. <u>La qualification du jury</u>	p98
3.2.2.2. <u>Les paramètres de l'échelle</u>	p99
3.2.2.3. <u>L'échantillon</u>	p100
3.2.3. Étude 1 : Une échelle perceptuelle pour l'étude du vibrato	p103
3.2.3.1. <u>Introduction</u>	p103
3.2.3.2. <u>Matériel et Méthode</u>	p110
3.2.3.3. <u>Résultats</u>	p113
3.2.3.4. <u>Discussion</u>	p121
3.2.3.5. <u>Conclusion</u>	p126

CHAPITRE IV : LES MESURES OBJECTIVES

4.1. Les mesures objectives manuelles	p129
4.1.1. La description des paramètres	p129
4.1.1.1. <u>Le temps maximum phonatoire</u>	p129
4.1.1.2. <u>La capacité vitale</u>	p130
4.1.1.3. <u>Le quotient phonatoire</u>	p131
4.1.1.4. <u>Le phonétogramme</u>	p132
4.2. L'apport des logiciels	p142
4.3. Les appareils employés dans notre recherche	p145
4.3.1. L'analyseur dynamique de signal Hewlett Packard	p145
4.3.1.1. <u>L'analyse spectrale</u>	p146
4.3.1.2. <u>Les applications de l'analyse spectrale</u>	p148
4.3.1.3. <u>Le rapport harmoniques/bruit</u>	p154

4.3.2 Le "Multi Dimensionnal Voice Program" (M.D.V.P.)	p157
4.3.2.1. <u>La description des paramètres</u>	p158
4.3.3. Les stations E.V.A. 1 et 2	p162
4.3.3.1. <u>La description des paramètres</u>	p164
4.3.3.2. <u>Les Mesures de la pression sous-glottique</u>	p168
4.3.3.3. <u>La pression sous-glottique estimée comme indice de forçage</u>	p175
4.3.3.4. <u>La pression sous-glottique estimée en fonction de l'âge</u>	p178
4.3.3.5. <u>La pression sous-glottique estimée en fonction de l'entraînement</u>	p181

CHAPITRE V : NOTRE CONTRIBUTION

5.1. Étude 2 : Application de l'analyse spectrale à la voix chantée

p186

5.1.1. Introduction	p186
5.1.2. Matériel et méthode	p187
5.1.3. Résultats	p187
5.1.4. Discussion et conclusion	p189

5.2. Étude 3 : Valeurs aérodynamiques et acoustiques de référence pour une population de sujets âgés de 60 à 90 ans

p193

5.2.1. Introduction	p193
5.2.2. Matériel et méthode	p193
5.2.3. Résultats	p195
5.2.4. Discussion et conclusion	p198

5.3. Étude 4 : Application du M.D.V.P. à la mesure du vibrato et corrélation avec une échelle perceptuelle
p201

5.3.1. Introduction	p201
5.3.2. Matériel et méthode	p203
5.3.3. Résultats	p207
5.3.4. Discussion et conclusion	p213

5.4. Étude 5 : Application de la station E.V.A. 1 à une population d'enfants âgés de 4 à 9 ans chanteurs et non-chanteurs
p216

5.4.1. Introduction	p216
5.4.2. Matériel et méthode	p216
5.4.3. Résultats	p218
5.4.4. Discussion et conclusion	p228

5.5. Étude 6 : Application de la station E.V.A.1 pour l'évaluation objective des dysphonies après paralysie d'une corde vocale.
P232

5.5.1. Introduction	p232
5.5.2. Matériel et méthode	p232
5.5.3. Résultats	p234
5.5.4. Discussion et conclusion	p237

5.6. Étude 7 : Application de la station E.V.A. 2 à une population d'adultes âgés de 20 à 50 ans pour l'obtention de valeurs de référence. Comparaison avec un groupe de sujets féminins dysphoniques
p243

5.6.1. Introduction	p243
---------------------	------

5.6.2. Matériel et méthode	p243
5.6.3. Résultats	p244
5.6.4. Discussion et conclusion	p258

CHAPITRE VI : INDEX DE LA QUALITÉ VOCALE

6.1. L'Index de Piccirillo	p264
6.2. Le "Dysphonia Severity Index" de Wuyts	p266

CHAPITRE VII : DISCUSSION GENERALE

CHAPITRE VIII : CONCLUSION

Bibliographie	p281
Tableaux	p309
Figures	P317
Annexes	p321

Introduction

La voix a l'insigne privilège d'assurer la communication entre les hommes ; que ce soit dans le quotidien ou dans des assemblées, par un cri spontané ou de longs discours complexes, par elle, ils transmettent leurs sentiments, leurs jugements, leurs pensées. Platon, lui-même, dans le "Phèdre" lui donnait le primauté sur le texte.

La production de la voix humaine met en mouvement des mécanismes contribuant non seulement à la production vocale mais aussi à la respiration et à la déglutition.

L'évaluation de la voix est multidimensionnelle. L'oreille humaine peut détecter des variations acoustiques subtiles et déterminer si ces variations sont de l'ordre du pathologique ou du normal. Cependant, l'audition est insuffisante pour fournir un diagnostic de base universel et guider la thérapie vocale ^[46, 52, 100, 327]. Le problème majeur de l'audition simple réside dans le fait qu'elle traite le signal acoustique dans sa globalité. L'oreille peut donc être dupée. Elle résiste difficilement au test de fiabilité intra et interindividuelle ^[171, 207, 208]. Elle est sujette aux variations psychologiques, à l'attention, à la fatigue, d'où l'intérêt de recourir à des instruments de mesures subjectives et objectives de la voix.

Les mesures subjectives qui reposent sur des échelles perceptuelles sont généralement effectuées par des médecins ou des spécialistes de la voix. Les mesures objectives telles que le temps maximum de phonation, le quotient phonatoire ou l'intensité existent depuis de nombreuses années ^[242]. Le développement de l'informatique et des logiciels a considérablement accru l'extension de leur champ, c'est ainsi que l'on peut mesurer les paramètres acoustiques tels que la fréquence, l'intensité vocale et leurs mesures de variabilité, de même que les paramètres aérodynamiques qui rapportent les mesures de flux aérien trans-glottique et de résistance de la glotte. Cependant, ces mesures, actuellement, ne sont pas standardisées, d'où la raison d'être de notre travail qui se propose de contribuer à la détermination de paramètres acoustiques et aérodynamiques qui entrent dans l'étude de la voix.

Pour ce faire, nous partons d'une revue de la littérature concernant les instruments d'évaluation déjà disponibles.

Un **premier chapitre** rappelle l'anatomophysiologie du larynx en nous limitant aux informations nécessaires pour comprendre le sujet.

Un **deuxième chapitre** aborde l'examen vidéostroboscopique pratiqué par l'oto-rhino-laryngologue (O.R.L.).

Un **troisième chapitre** développe les échelles d'auto-évaluations qui concernent l'avis du patient quant à l'estimation de sa qualité vocale. Ce chapitre rend compte des diverses mesures perceptuelles rencontrées dans la littérature et mises à la disposition des cliniciens pour évaluer la qualité vocale du patient. Nous y détaillons plus amplement l'échelle G.R.B.A.S.^[122] (en anglais, Grade, Rough, , Breath, Astheny, Strain soit grade, raucité, souffle, sensation de forçage, sensation de faiblesse) qui satisfait de nombreux praticiens. Nous apportons notre contribution personnelle en proposant une échelle perceptuelle. Nous avons proposé une échelle perceptuelle composée de six paramètres destinés à la voix chantée et plus spécifiquement à l'analyse du vibrato (étude 1). Les quatre questions suivantes ont guidé notre travail.

- Un haut niveau d'expertise renforce-t-il la concordance des jugements lors de l'évaluation du vibrato ?
- Quelle est la reproductibilité intra et inter-juges du choix catégoriel (vibrato, son blanc, trémolo, chevrotement, autre) ?
- Quelle est la reproductibilité de l'échelle-test portant sur les paramètres non-génant, harmonieux, agréable, régulier, stable et timbré ?
- Quel est l'impact de la complexité de la tâche choisie (morceau imposé versus voyelle isolée) sur l'évaluation perceptuelle de l'oscillation ?

Le **quatrième chapitre** présente les mesures acoustiques et aérodynamiques. Nous y exposons les logiciels mis à notre disposition, soit l'analyseur spectral de Hewlett Packard^[117], le Computer Speech lab. de Kay Elemetrics^[159, 160] et plus spécifiquement son programme M.D.V.P. (Multi-Dimensionnal Voice Program)^[159, 160] ainsi que les stations E.V.A. 1 et 2 (Évaluation Vocale Assistée) mises au point par le Laboratoire Parole et Langage de Marseille^[293, 296].

Le **cinquième chapitre** rend compte de nos études, à savoir :

1. Les travaux avec l'analyseur spectral (Étude 2) sur un groupe de trente chanteurs qui nous ont permis d'observer les modifications de la courbe spectrale en voix chantée.
2. L'objectivation de la qualité vocale selon le même procédé pour une population de sujets âgés de 70 à 90 ans et la comparaison de nos résultats aux mesures aérodynamiques à ceux d'un groupe de sujets jeunes (Étude 3).
3. L'élaboration d'un lien éventuel entre les mesures subjectives (jury d'écoute) et objectives (par le M.D.V.P. et la station E.V.A. 1) à partir de deux études, l'une portant sur un groupe de chanteurs (Étude 4), et l'autre sur un groupe de sujets atteints de paralysie unilatérale laryngée (Étude 6).
4. L'établissement de valeurs de référence pour les paramètres acoustiques de la station E.V.A. 1 pour des enfants âgés de 4 à 9 ans répartis en un groupes de chanteurs et un groupe de non-chanteurs mesurant ainsi l'effet de l'entraînement vocal (Étude 5).
5. L'établissement de références acoustiques et aérodynamiques à partir de la station E.V.A. 2 sur une population de 88 adultes âgés de 20 à 50 ans, de la mesure de la pression sous-glottique estimée et de sa comparaison à un groupe de 15 patients porteurs de nodules ou d'œdème de Reinke (Étude 7).

Le **sixième chapitre** introduit la notion d'index de la qualité vocale. Nous n'apportons ici aucune contribution personnelle. Le but de l'index est de réduire les mesures acoustiques et aérodynamiques à une seule valeur. Nous présentons l'index de sévérité de la dysphonie (D.S.I. – Dysphonia Severity Index ^[356]) développé par le groupe belge d'étude de la voix (B.S.G.V.D.¹)^[332] et auquel nous avons participé.

Le **septième chapitre** comporte la discussion générale et le **huitième** la conclusion de l'ensemble de notre travail.

¹ *B.S.G.V.D.* : abréviation de *Belgium Study Group for Voice disorders*. Ce groupe a commencé ces travaux en 1994 et regroupe différents centres : l'hôpital universitaire d'Anvers (UZA), de Bruxelles (VUB), de Gent (UZG), de Louvain Mont-Godinne (UCL), du service logopédique de Turnhout (KMSL) et le cabinet O.R.L. du Dr B. Millet, (Van de Heyning PH, Remacle M, Van Cauwenberghe (1996)^[332]).

Chapitre 7

L'anatomophysiologie
des organes phonateurs

1.1. L'ANATOMIE

1.1.1. Les organes de la phonation.

Le larynx est essentiel à la production du son chez le sujet normal. Cependant comme le souligne Sataloff ^[267], l'anatomie vocale ne se limite pas au larynx. Le mécanisme vocal englobe la musculature abdominale et dorsale, la cage thoracique, les poumons et le pharynx, la cavité orale et le nez. Chacune de ces composantes assure une fonction importante dans la production vocale. La position et les mouvements corporels jouent un rôle non négligeable sur la phonation. Ils peuvent être responsables d'un dysfonctionnement vocal ^[276].

Le tractus vocal contient essentiellement le **boîtier ostéo-cartilagineux** laryngé composé des cartilages thyroïde, cricoïde et aryténoïdes, de l'os hyoïde et des corniculés [figure 1]. Il s'articule grâce aux muscles intrinsèques et extrinsèques qui lui confèrent une grande mobilité. Les cordes vocales ancrées au sein du cartilage thyroïde constitue l'espace glottique. Au-dessus et en-dessous de cet espace se situent respectivement les tractus vocaux supra- et sous-glottiques. Les bandes ventriculaires ou fausses cordes sont localisées au-dessus des cordes vocales. Elles n'entrent pas en contact durant la phonation ou le chant.

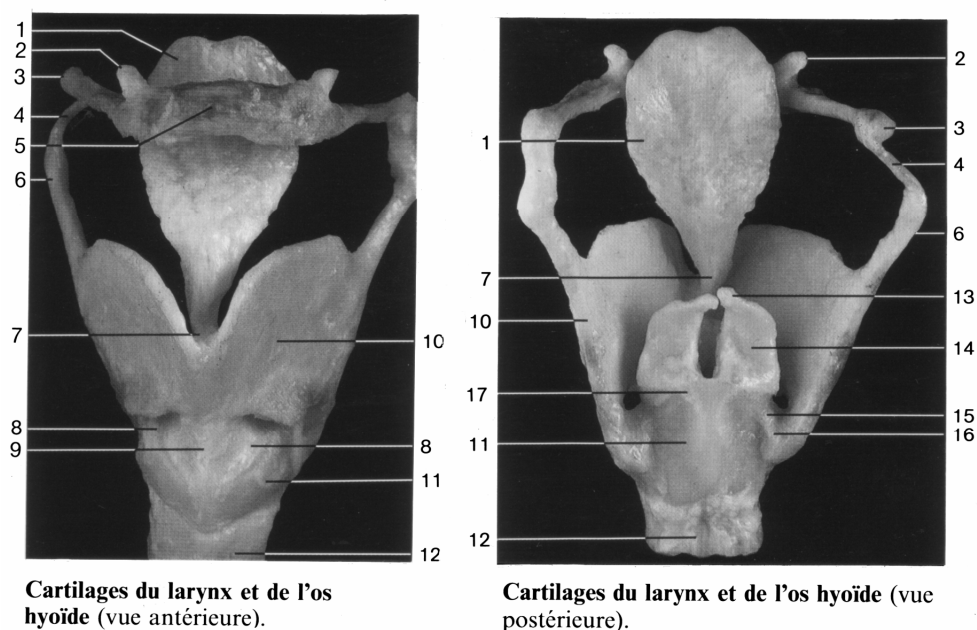


Figure 1 : Boîtier ostéocartilagineux.

1.1.2. Les ligaments

Les **ligaments** sont des membranes de tissus conjonctifs et élastiques unissant les différents constituants squelettiques du larynx. Les ligaments extrinsèques sont les ligaments thyro-hyoïdiens, et la membrane crico-trachéale. Les ligaments intrinsèques sont la membrane crico-thyroïdienne, le ligament thyro-épiglottique, les ligaments crico-aryténoïdiens, le ligament crico-corniculé, la membrane fibro-élastique, tendue entre l'épiglotte et les cartilages aryténoïdes, constituant ainsi les ligaments aryténo-épiglottiques. Cette membrane se prolonge sous la muqueuse et présente de chaque côté deux épaississements : le ligament vestibulaire et le ligament vocal [figure 2].

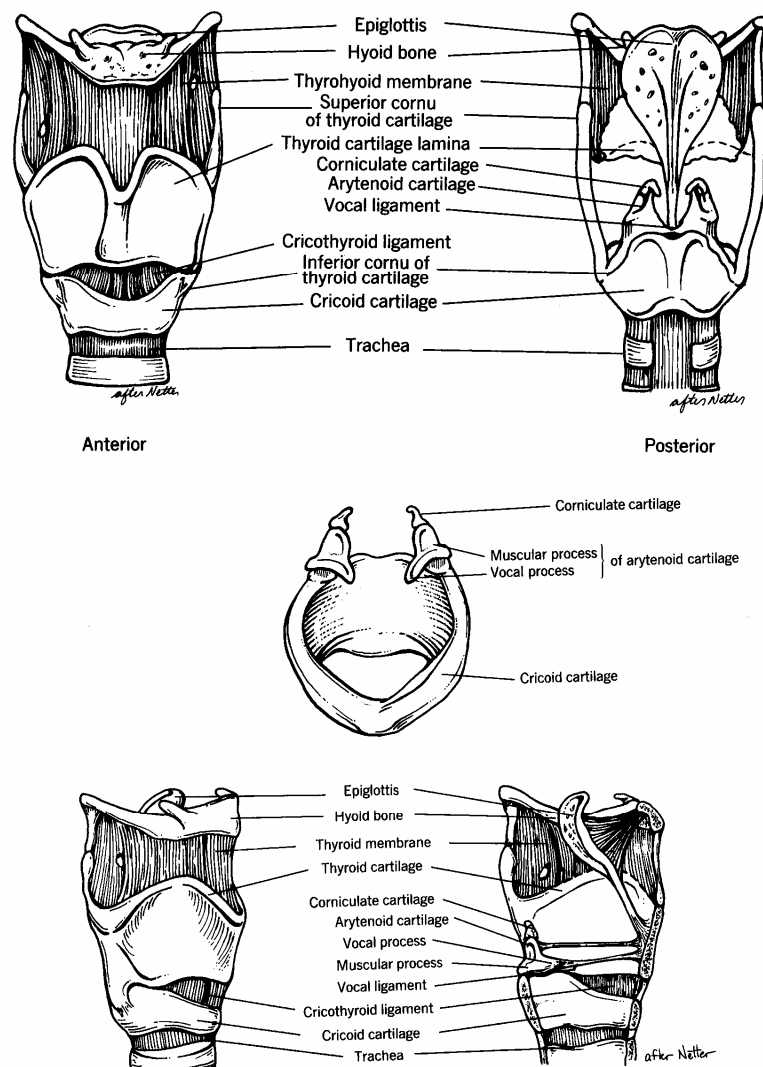


Figure 2 : Cartilages, membranes et ligaments du larynx.

1.1.3. Les muscles intrinsèques

Les muscles intrinsèques du larynx sont reliés aux cartilages précités [figure 3].

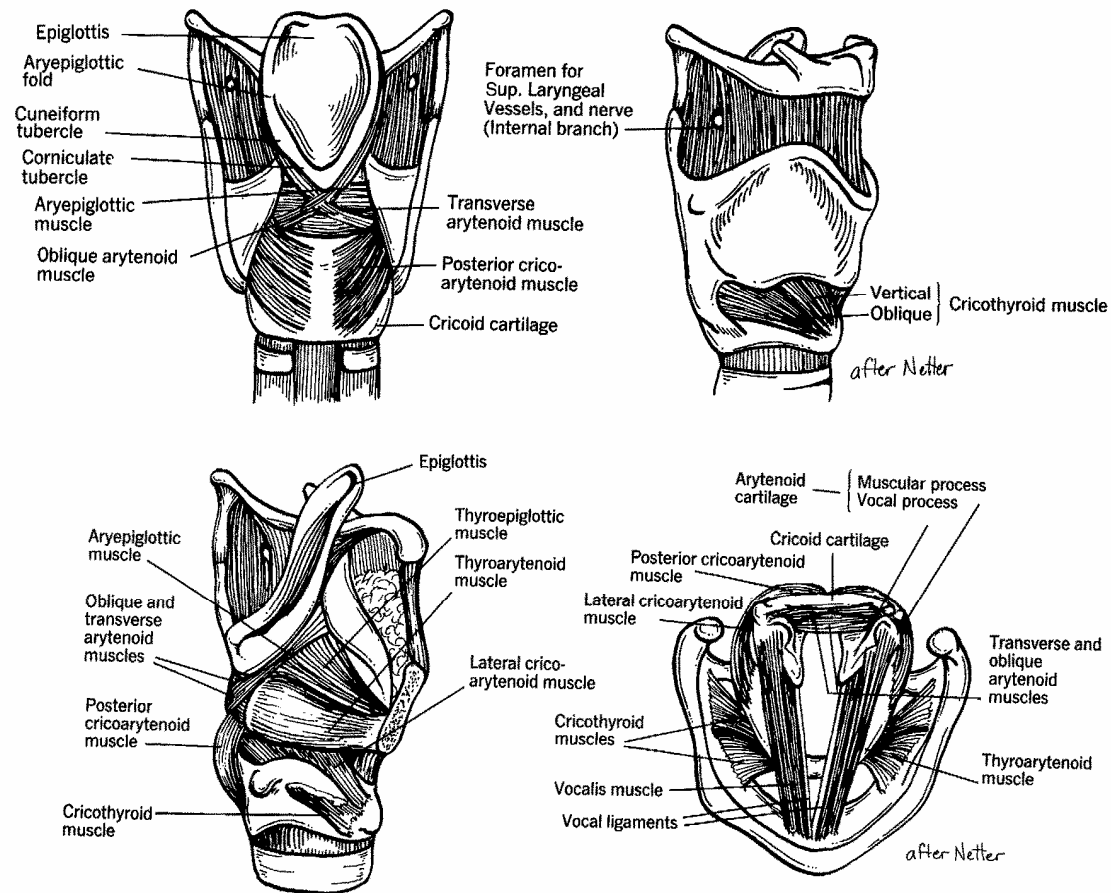


Figure 3 : Muscles intrinsèques du larynx.

Le muscle thyro-aryténoïdien inférieur du larynx s'étend sur chaque côté des cartilages aryténoïdiens à l'intérieur du cartilage thyroïde juste en-dessous et derrière la proéminence de la pomme d'Adam. Il forme le corps des cordes vocales. Les cordes vocales agissent comme un vibreur, la source sonore du tractus vocal.

Les muscles intrinsèques modifient la position, la forme, et la tension des cordes vocales. Ils jouent un rôle dans leur rapprochement (adduction), leur éloignement (abduction), ou leur élongation en diminuant la tension longitudinale. Les aryténoïdes permettent un mouvement complexe des cordes vocales par leur mobilité (balancement, glissement, rotation). Ils modifient ainsi la forme du bord libre des cordes vocales. Tous les muscles de part et d'autre du larynx (excepté le muscle

crico-thyroïdien) sont innervés par le nerf laryngé inférieur appelé aussi récurrent [Figure 4].

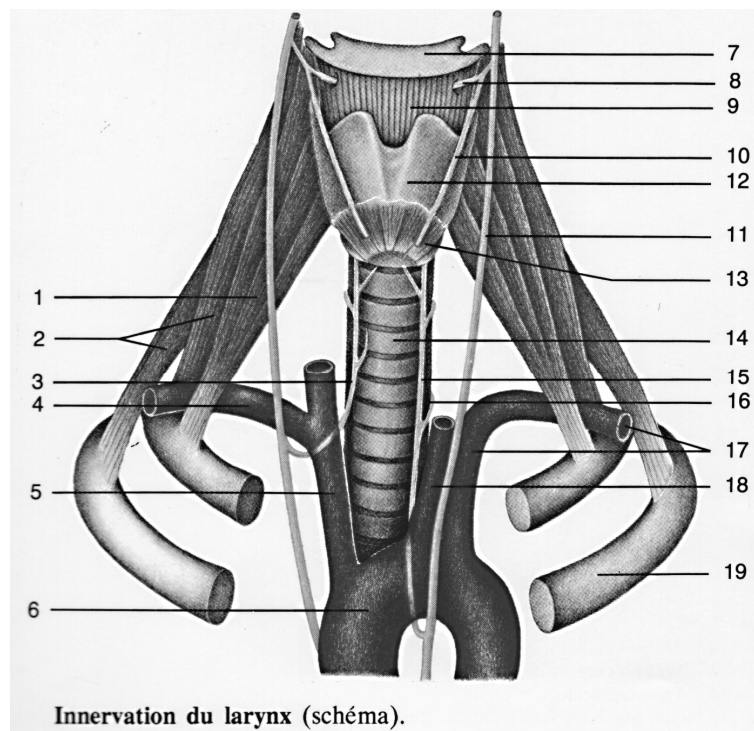


Figure 4 : Innervation du larynx.

Ils sont au nombre de quinze. Excepté le muscle inter-aryténoïdien, ils vont par paire soit les crico-thyroïdiens, les crico-aryténoïdiens postérieurs et les latéraux, les thyro-aryténoïdiens supérieurs et les inférieurs (les muscles vocaux), l'aryténoïdien oblique, les ary-épiglottiques.

Ces structures soutiennent le larynx. Le muscle crico-thyroïdien est innervé par le nerf laryngé supérieur sur chaque côté. La tension des cordes vocales devient possible grâce au contrôle des muscles crico-thyroïdiens. Ces derniers naissent de la face latérale de l'arc cricoïdien et se clivent en deux faisceaux : un faisceau oblique et un faisceau horizontal. Les structures des cordes vocales se modifient en fonction de leur degré d'élongation (spécialement de la relation entre la couche superficielle du muscle vocal, la couverture et le muscle vocal profond, le corps). Sonninen et al (1998)^[289] rapportent que selon Alipour et al (1999)^[3], l'élongation des cordes vocales

est de maximum 20 % en phonation. Pour une élongation de 0 à 20%, , la tension dans le corps est principalement sous le contrôle du muscle vocal. Pour une élongation supérieure à 20%, les muscles extrinsèques des cordes vocales contrôlent la tension du muscle vocal profond. Selon Vilkmann cité par Alipour et al (1999)^[3], 30 à 40% de l'élongation des cordes vocales sont dus au glissement et à la rotation du cartilage cricoïde.

1.1.4. Les muscles extrinsèques.

- Les muscles extrinsèques relient le larynx à la mandibule, à l'os hyoïde ainsi qu'au sternum et aux clavicules. Ils permettent au larynx des mouvements dans le plan vertical qui assurent une mobilité importante pour la déglutition mais aussi durant le chant.

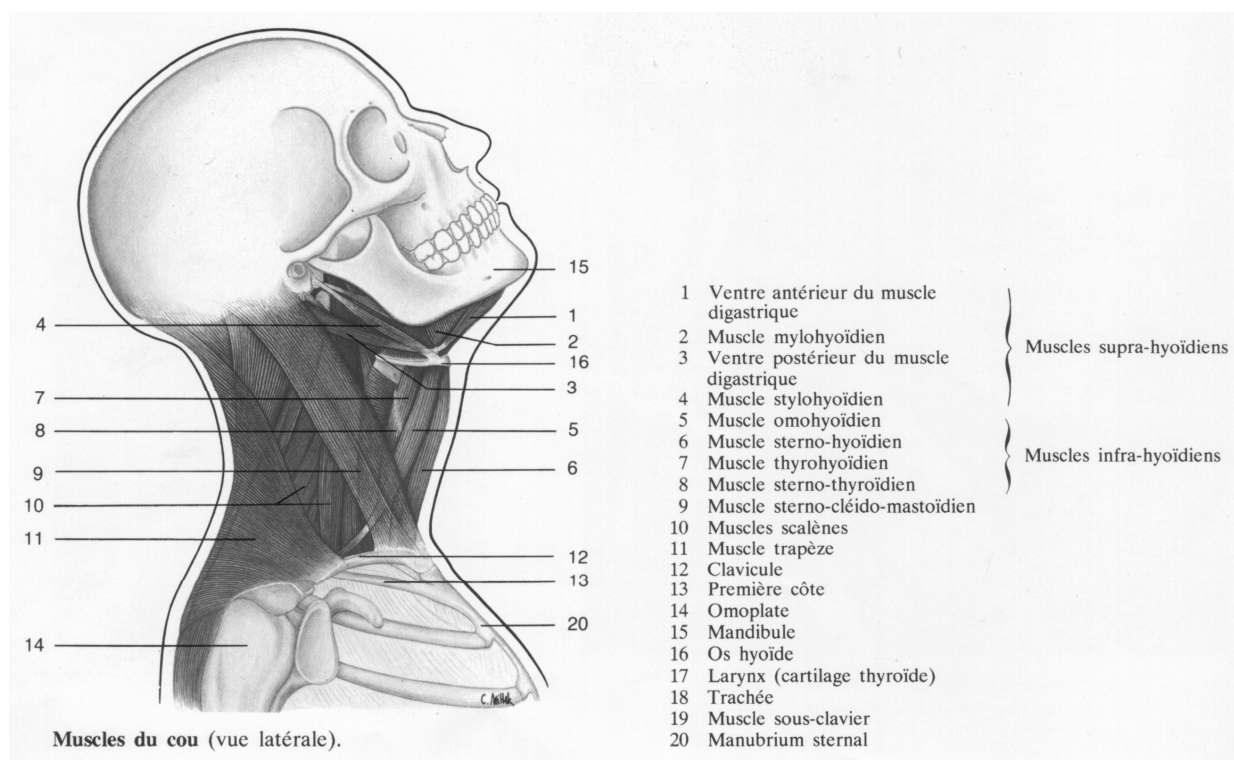


Figure 5 : Muscles extrinsèques (vue latérale).

On les répartit en deux groupes.

- Les muscles sous-hyoïdiens rattachant le larynx à la partie supérieure du thorax ;
- Les muscles sus-hyoïdiens reliant le larynx à la base du crâne et à la mâchoire inférieure.

Les muscles éleveurs du larynx sont les muscles digastriques, les stylo-hyoïdiens, les stylo-pharyngiens, les génio-hyoïdiens, les mylo-hyoïdiens, les génio-glosses, les myo-glosses, le constricteur moyen du pharynx.

Les muscles abaisseurs du larynx sont les muscles sterno-thyroïdiens, les thyro-hyoïdiens, les sterno-hyoïdiens, les omo-hyoïdiens. [Figure 5]

1.1.5. Structure de la corde vocale

Selon Hirano et al (1983)^[123], la corde vocale est composée de 5 couches fines [Figure 6]

L'épithélium de type épidermoïde pavimenteux stratifié non kératinisant représente une surface de contact avec la partie vibrante des cordes vocales. Il est bien adapté pour résister au traumatisme occasionné par le contact des cordes vocales. Il encapsule le tissu plus fluide de la sous-muqueuse appelée lamina propria et constituée de trois couches^[98].

La **couche superficielle** d'une épaisseur de 0.5 mm se compose d'un enchevêtrement de fibres d'élastine peu organisées dans une sorte de fluide interstitiel, c'est l'espace de Reinke. Ces fibres d'élastine ont une structure protéique adaptée aux contraintes de l'étirement longitudinal.

La **couche intermédiaire** est également constituée de fibres d'élastine et de fibres de collagène, les fibres d'élastine étant préférentiellement orientées dans le sens antéro-postérieur.

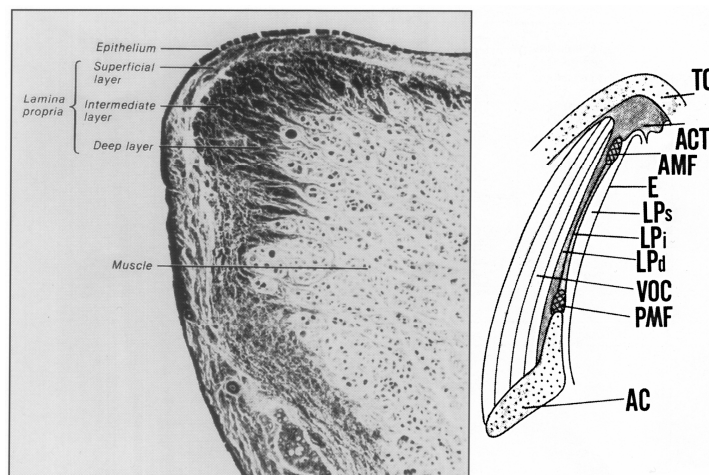
La **couche profonde** est surtout constituée de fibres de collagène enchevêtrées. Cette structure protéique est moins favorable à l'élongation que celle des fibres d'élastine.

Les **couches intermédiaire et profonde** de la lamina propria forment le ligament vocal. Son épaisseur varie entre 1 et 2 mm. Dans cette couche intermédiaire, les extrémités de la corde vocale appelées macula flava antérieure et macula flava postérieure sont faites d'une masse ovale épaisse. Ces deux masses sont composées de fines mailles de fibres élastiques, de fibroblastes et de tissus conjonctifs de soutien. Dans la partie antérieure de la macula flava se situe le tendon de la commissure antérieure essentiellement composée de collagène. Au niveau postéro-latéral, il adhère à la couche profonde de la lamina propria et est ainsi en contact avec la macula flava antérieure. Dans sa partie antérieure, il rejoint le cartilage thyroïde.

La macula flava postérieure est rattachée à l'apophyse vocale de l'aryténoïde par une petite surface de transition constituée de chondrocytes, de fibroblastes et de cellules non identifiées dont la constitution se rapproche des deux autres types de cellules.

Ces structures renforcent et protègent les extrémités du ligament vocal des éventuelles blessures mécaniques causées par la vibration des cordes vocales.

Le muscle thyro-aryténoïdien inférieur constitue la masse musculaire de la corde vocale. Son épaisseur est de 7 à 8 mm au niveau du tiers moyen.



Légende : TC : cartilage thyroïde, ACT : tendon de la commissure antérieure, AMF : macula flava antérieure, PMF : macula flava postérieure, AC : cartilage aryténoïde, E : épithélium, LP : lamina propria, LPs: lamina propria couche superficielle, Lpi : lamina propria couche intermédiaire, Lpd : lamina propria couche profonde, Voc : muscle vocal.

Figure 6 : Couches de la corde vocale.

Le muscle vocal est divisé en deux : le vocalis superficiel et le muscularis profond. Des hypothèses sont élaborées quant à leur fonction. Le muscularis serait utilisé pour des raccourcissements rapides de la corde vocale, tandis que le vocalis

pourrait être utilisé pour le réglage fin de la tension des cordes (Giovanni (1998)^[98]). Ces observations sont d'un grand intérêt dans l'apprentissage du chant et la compréhension des ajustements vocaux adéquats des chanteurs.

Les vaisseaux sanguins sont situés dans la muqueuse du bord libre de la corde vocale. Ils proviennent à la fois de l'extrémité antérieure et postérieure de la corde vocale membraneuse et en longent le bord. Quelques vaisseaux, très minces, proviennent directement du muscle vocal et pénètrent dans la muqueuse. Cette configuration vasculaire favorise le maintien de la vibration.

1.1.6. Le tractus vocal supra-glottique

Baken (1987)^[15] compare le tractus supra-glottique à un tuyau d'orgue. Il se comporte comme un tube résonnant, incluant le pharynx, la langue, le palais, la cavité orale, le nez. Ensemble, ils agissent comme des résonateurs. Ils sont responsables de la qualité du timbre vocal, et du caractère perçu de tous les sons parlés. La vibration des cordes vocales à la source produit un son de bourdonnement "buzzing" (Sundberg (1977)^[305]).

1.1.7. Le tractus vocal sous-glottique

Le tractus sous-glottique canalise l'air en provenance des poumons. Les chanteurs ou les acteurs se réfèrent à cette source vocale et en parlent en terme de soutien diaphragmatique. Des déficiences dans le soutien sont fréquemment responsables de la dysfonction vocale. Ce soutien mécanique produit une force qui dirige un flux d'air contrôlé entre les cordes vocales. Les muscles respiratoires actifs travaillent de concert avec la force de résistance passive de l'espace glottique.

La phonation exige une interaction entre la source d'énergie, le vibreur, et les résonateurs. Le fonctionnement laryngé qui permet l'émission laryngée et les caractéristiques vocales qui en découlent sont développées aux points suivants.

1.2. LA PHYSIOLOGIE.

1.2.1. Le Mécanisme vibratoire des cordes vocales.

Les théories prises en considération sont :

1. **La théorie aérodynamique myo-élastique**

a. La théorie myo élastique d'Ewald (1898) :

La vibration cordale est un mouvement passif résultant de la rupture de l'équilibre entre la tension des cordes et la pression sous-glottique. En effet, l'espace glottique rétréci par la contraction des muscles tenseurs et adducteurs des cordes vocales oppose une résistance au passage de l'air provenant des poumons. La pression de la colonne d'air augmente avec l'expiration, ce qui entraîne un écartement des cordes vocales, une fuite d'air, une chute de la pression sous- glottique et ensuite un nouvel accollement des cordes puisque la tension du muscle vocal est maintenue et que la structure élastique le permet.

b. La théorie aérodynamique myo-élastique complétée (1962) :

Cette théorie est basée sur deux forces opposées qui sont la pression sous-glottique et la tension musculaire des cordes vocales. L'effet Bernouilli² induit par la forme anatomique des cavités supra-laryngées permet la vibration des cordes vocales avant qu'elles ne soient complètement rapprochées. Cette opposition favorise leur écartement et leur rapprochement par un effet d'aspiration.

Pendant la phonation, la glotte s'ouvre et se ferme produisant ainsi des changements de forme et de résistance. Pour répondre à ces modifications, des réajustements rapides et complexes de la musculature supra-glottique sont nécessaires. Sous l'effet de la pression sous-glottique, les cordes vocales s'écartent progressivement pour laisser s'écouler l'air. Dès que cette quantité d'air s'est échappée, les bords libres se rapprochent sous l'effet de la diminution de la pression sous-glottique, des caractéristiques mécaniques des cordes vocales, et du

² *L'Effet Bernouilli : après le passage dans un tube non rigide d'un flux laminaire à une vitesse minimale, la pression négative qui en résulte, tend à aspirer les parois, dans ce cas la muqueuse des cordes cordes vocales.*

phénomène Bernouilli. La partie inférieure de la glotte commence à se fermer alors que les bords supérieurs sont toujours en train de se séparer. La partie supérieure des cordes vocales possèdent des propriétés élastiques très fortes, ce qui les ramènent très rapidement sur la ligne médiane. Les cordes vocales sont de nouveau en contact. La pression pulmonaire étant toujours active, le même phénomène d'écartement va se reproduire périodiquement. La pression sous-glottique commence à se reformer et l'événement se répète.

Ce déplacement en vague des cordes vocales produit une ondulation de muqueuse qui peut être examinée cliniquement sous la lumière stroboscopique, [Figure 7].

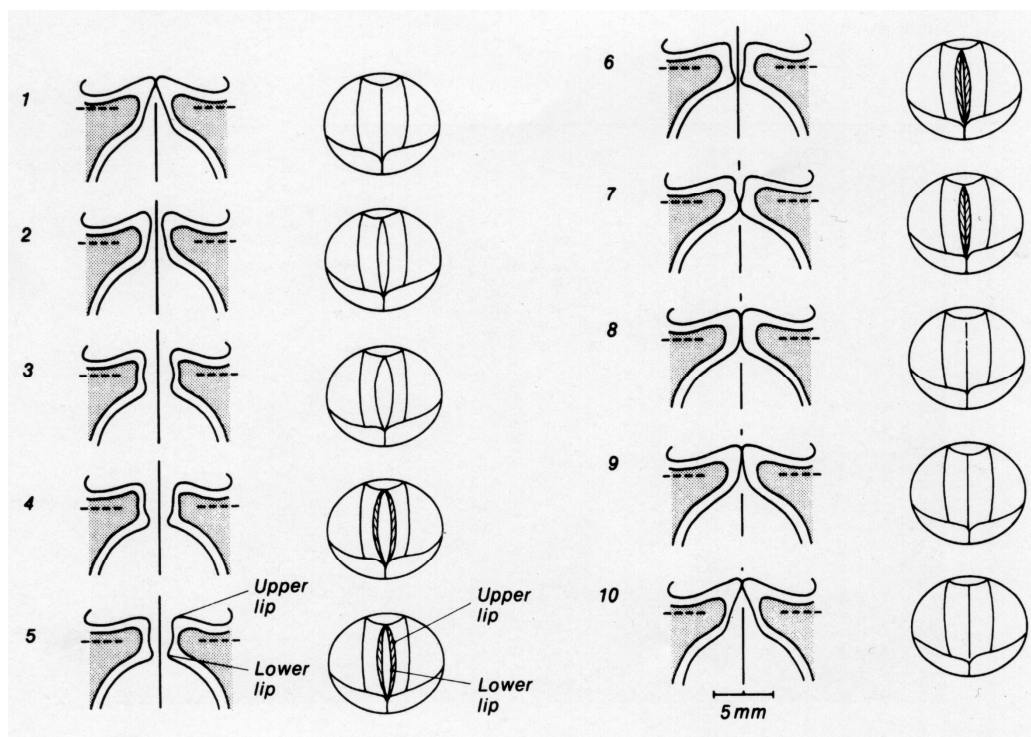


Figure 7 : Vibration des cordes vocales en coupe frontale.

2. La théorie oscillo-impédantielle (Dejonckere (1985)^[57]) complète finement les concepts précédents. L'auteur conçoit la corde vocale en vibration comme un oscillateur entretenu. La fonction de ce vibreur est déterminée par l'inertie et l'élasticité de la partie vibrante. L'onde de pression sous-glottique constitue son apport périodique d'énergie. L'auteur a observé que les cordes vocales durant la phase terminale de l'émission sonore, vibraient avec une amplitude décroissante avant de regagner leur position initiale. Le larynx serait assimilable à un vibreur à faible degré d'amortissement. Il n'engendre pas de vibrations au niveau des cordes vocales. Ces

vibrations proviendraient bien de la pression sous-glottique. La structure des fibres musculaires de la corde vocale lui permet d'obtenir activement une forme légèrement concave 300 millisecondes avant le début de l'émission sonore. La glotte présente alors une forme de fuseau permettant l'apparition de turbulences et la mise en route de l'effet Bernouilli. Le flux aérien, d'abord continu, amorcerait la vibration cordale à son passage.

Dejonckere décrit des sous-vibrateurs :

- l'un à composante horizontale soit l'amplitude d'écartement de surface des cordes vocales ,
- l'autre verticale soit le soulèvement du bord de la corde vocale à chaque oscillation,
- le troisième, constitué par le passage de l'ondulation de muqueuse.

L'impulsion de la première oscillation résulterait du déséquilibre entre la pression sous-glottique et la tension de fermeture des cordes vocales. La pression est nécessaire à la mise en vibration de la corde vocale. Elle augmente avec la viscosité de la corde, l'importance de l'espace glottique pré-phonatoire et la tension des cordes vocales (Dejonckere 1995^[63]).

L'impédance glottique est dépendante de l'état visco-élastique de la muqueuse, et de la régulation tonique et kinétique des muscles intrinsèques. Elle est définie par cette équation :

$$\text{Impédance} = \frac{\text{Pression sous-glottique} - \text{Pression supra-glottique}}{\text{Débit trans-glottique}}$$

La pression sous-glottique croît jusqu'à plus ou moins 7cm/H₂O pour le langage courant. La fréquence de la vibration ou hauteur tonale (soit le nombre de cycles d'ouverture et de fermeture par seconde mesuré en Hertz) dépend des propriétés mécaniques et de la pression d'air appliquées sur les cordes vocales, qui sont conditionnées par les muscles laryngés.

1.2.2. Caractéristiques acoustiques de la voix.

La hauteur, l'intensité et le timbre sont des caractéristiques particulièrement importantes pour le contrôle de la voix. Nous en donnons ici les définitions.

1.2.2.1. Définitions

La hauteur fut à l'origine un terme musical. Elle devint par la suite un terme psychologique servant à désigner une qualité de son perçue. En musique, la hauteur des sons a été perçue bien avant que les bases physiques ne soient comprises. L'une des grandes découvertes musicales est que pour les sons périodiques, comme ceux produits par la voix humaine, la hauteur est inévitablement liée à la périodicité de l'onde sonore (Pierce J (1984)^[233]). Elle est évaluée en hertz (Hz).

L'intensité est proportionnelle à l'amplitude du mouvement vibratoire des cordes vocales. Elle dépend essentiellement de la valeur de la pression sous-glottique qui elle-même résulte de la coordination entre la force du soutien abdominal et la force de fermeture des cordes vocales. Elle est évaluée en décibel (dB). En voix conversationnelle, le niveau est d'à peu près 60 dB, mais il peut varier de 40 dB pour la voix chuchotée à 120 dB pour les voix d'opéra.

Le timbre est fonction de la quantité d'harmoniques du son, de la modulation du son initial à travers les cavités de résonance (pharynx, cavité buccale, et fosses nasales). On distingue le timbre vocalique et le timbre extra-vocalique. Le timbre vocalique, commun à tous les individus, correspond aux zones formantiques qui permettent de reconnaître la hauteur d'une voyelle au-delà des variabilités d'expression individuelles. Le timbre extra-vocalique détermine les caractères personnels de la voix. C'est en quelque sorte une carte d'identité vocale. Sa couleur permet d'identifier la personne au même titre que son visage ou sa silhouette.

2.2.2.2. La hauteur

La fréquence fondamentale, ou plus simplement le fondamental qui correspond à la hauteur tonale, peut être modifiée par les propriétés mécaniques des cordes vocales (tension et rigidité) et par la pression d'air (Titze et al (1994)^[327]). Lorsque les cordes vocales sont amincies et étirées, et que la pression d'air est augmentée, la fréquence de l'émission des pulsations d'air et l'intensité augmentent. Pour comprendre ce mécanisme myo-élastique et aérodynamique à la base de la phonation, il est important de noter que les cordes vocales laissent échapper de petites giclées d'air plutôt que d'induire une réelle mise en vibration des cordes vocales. Il y a une différence dans la phase verticale. La partie inférieure des cordes vocales commence à ouvrir et à fermer la glotte avant la partie supérieure [cfr figure 7 ci-dessus].

Le fondamental est donc indicatif des caractéristiques biomécaniques des cordes vocales lorsqu'elles entrent en interaction avec la pression sous-glottique. Ces propriétés sont décrites par la structure laryngée et les forces musculaires qui y sont appliquées. L'ajustement de ces forces est fonction de comportements réflexes, ou d'un enseignement adéquat (Miller (1986)^[200]). La capacité de pouvoir varier le fondamental renseigne sur l'efficacité mécanique des structures laryngées, sur la précision de leur contrôle. Le degré de perturbation en fréquence fournit un indice de la stabilité du système phonatoire.

La contraction du muscle crico-thyroïdien fait pivoter le cartilage thyroïde et augmente ainsi la distance entre le thyroïde et les cartilages aryénoïdes. Plus exactement, la partie verticale (pars recta) du crico-thyroïdien élève le cricoïde et la partie oblique (pars obliqua) entraîne une rétraction du cartilage cricoïde et un glissement postérieur du cartilage thyroïde. En conséquence, les cordes vocales sont allongées. Ce mouvement de bascule augmente la surface exposée à la pression sous-glottique et rend la pression d'air plus efficace dans l'axe d'ouverture de la glotte. De plus les fibres élastiques des cordes vocales étirées leur permettent de rejoindre facilement la ligne médiane. Les cycles sont raccourcis et se répètent plus souvent. La fréquence fondamentale augmente [Figure 8].

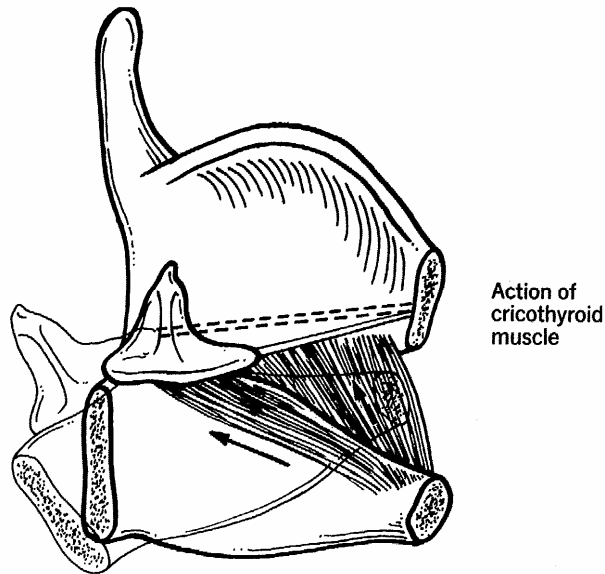


Figure 8 : Mouvement de bascule du cartilage thyroïde sur le cartilage cricoïde par l'action du muscle thyro-cricoïdien.

Le temps de réponse du muscle aiderait à déterminer la vitesse à laquelle les changements de hauteur tonale et de registre peuvent être réalisés durant la phonation. Pour les changements de fréquence, le crico-thyroïdien agit probablement dans un mode isotonique, changeant sa longueur durant la contraction. La corde vocale s'accroît de 0.7 à 3.5 mm par octave (Nishizawa (1988) in Alipour et al (1999)^[3]). Sonninen et al (1999)^[290], par des méthodes radiographiques, étudie le contrôle de la hauteur tonale, le registre et le mode de chant en fonction des mouvements des structures laryngo-pharyngées. Il observe que la hauteur tonale liée à la distance du thyro-aryténoïdien croît en proportion inverse de cette distance et qu'il y a donc une réduction de la distance dans les hautes fréquences, ce qui suggère l'emploi d'un mécanisme alternatif pour le contrôle de la hauteur. Alipour et al (1999)^[3] ont montré que le muscle crico-thyroïdien peut changer de longueur. Une diminution de plus de 50 % de la longueur du crico-thyroïdien a été observée pour les hauteurs tonales élevées. La vitesse maximale de changement de la hauteur tonale est de l'ordre de 60 à 100 millisecondes et une modification de 50% de la longueur du muscle serait réalisée en 200 millisecondes. Les auteurs nuancent leur résultat. Dans la réalité, la contraction du thyro-aryténoïdien s'oppose à cette modification de longueur.

Titze et al (1989)^[323] expliquent l'importance des propriétés élastiques. Si le muscle crico-thyroïdien se contracte tandis que le thyro-aryténoïdien est inactif, il en résulte un allongement de la corde vocale, une rigidité des tissus et une élévation de la fréquence fondamentale. Si par contre le muscle thyro-aryténoïdien se contracte et le crico-thyroïdien est passif, il s'ensuit un raccourcissement de la corde vocale, une diminution de la rigidité de la couverture, et la possibilité que la rigidité du corps augmente. Dès lors l'épaisseur de la corde pourrait augmenter deux fois plus vite. En cas de paralysie du thyro-aryténoïdien, les modifications de fréquence seraient dépendantes du crico-thyroïdien, aussi le temps de réponse pourrait considérablement ralentir. Le temps de contraction tétanique du crico-thyroïdien rapporté par Alipour et al (1999)^[3] est de plus ou moins 192 millisecondes pour une contraction tétanique de 90 % et 117 millisecondes pour le thyro-aryténoïdien.

Deux modes vibratoires différents existent en fonction du type d'émission fréquentiel laryngé, sons graves ou sons aigus. L'émission de sons graves se produit en mécanisme 1 dit lourd, celle de sons aigus en mécanisme 2 dit léger ^[269].

En mécanisme de type 1, les muscles ary-aryténoïdiens et crico-aryténoïdiens latéraux sont contractés. Ils mettent les cordes vocales en position de phonation. Le muscle thyro-aryténoïdien inférieur dit muscle vocal est contracté, ce qui provoque l'épaississement de la surface de contact du bord libre des cordes. Il vibre sur toute son épaisseur. Les cordes vocales sont courtes. Le contact entre les deux bords libres se fait sur une profondeur de plusieurs millimètres. La stroboscopie permet d'observer une ondulation de muqueuse ample. Davantage de tension des muscles vocaux sont nécessaires pour accroître la fréquence vocale. Le degré de pression d'air sous-glottique accroît l'intensité. Il s'agit d'un équilibre entre la pression sous-glottique, la pression supra-glottique et la pression de fermeture. Cet équilibre évite de mettre les cordes vocales sous trop forte tension.

En mécanisme de type 2, les muscles ary-aryténoïdiens et crico-aryténoïdiens latéraux augmentent leur contraction, ce qui entraîne une augmentation de la compression médiane des aryténoïdes. Les muscles vocaux sont détendus ; par contre, les crico-thyroïdiens se contractent de plus en plus, provoquant un étirement

des cordes vocales qui s'amincissent : seul le bord libre va vibrer. La surface d'accolement est très petite. En stroboscopie, l'amplitude vibratoire diminue. L'ondulation de la muqueuse est réduite, le temps de fermeture raccourci. Dans les émissions les plus aiguës, il n'y a plus d'accolement. Quand la pression sous-glottique augmente, la résistance des ligaments vocaux s'accroît modérément. Dans ce mode de vibration, on joue essentiellement avec l'élasticité de la glotte. Les muscles vocaux perdent leur rôle prédominant.

L'augmentation de fréquence se fait par l'élévation de la résistance glottique, par l'augmentation de tension élastique du ligament vocal et par le raccourcissement de la partie vibrante.

Le passage d'un mécanisme à l'autre débute par un affaiblissement du timbre et de l'intensité. Il se situe entre 258 et 345 Hz chez l'homme et 274 et 318 Hz chez la femme. Cette zone de transition est perceptible dans le timbre. A l'intérieur de cet intervalle, les mécanismes se recouvrent.

1.2.2.2.1. Les facteurs qui affectent la hauteur.

Des auteurs, dans la littérature, tels que Fitch et al (1990)^[81], Horii et al (1981)^[139], Hudson et al (1982)^[141], Snidecor (1951)^[287], Vuorenkoski et al (1978)^[338], Abitbol et al (1999)^[2]..., mentionnent l'âge, la pathologie, la culture, l'état émotionnel, les facteurs hormonaux, le sexe, la posture comme éléments principaux influençant la fréquence fondamentale du locuteur et du chanteur.

1.2.2.2.1.1. L'âge.

Sataloff (1997)^[269] insiste sur la bonne compréhension de l'évolution et du développement de la voix. Comprendre les différences anatomophysiologiques entre les enfants, les adolescents, les adultes et les personnes âgées permet tant aux thérapeutes qu'aux enseignants de dispenser un entraînement optimal avec un risque minime de lésion vocale. Il rappelle les données suivantes : la longueur des cordes vocales dans l'enfance est de 6 à 8 mm pour atteindre 17 à 23 mm à l'âge adulte. L'ossification des cartilages du larynx s'effectue progressivement en fonction de l'âge.

A 2 ans l'os hyoïde est ossifié. A 20 ans le cartilage thyroïde et cricoïde s'ossifient, à 30 ans les aryténoïdes, et après 65 ans les cartilages cunéiformes et corniculés. La fréquence fondamentale à la naissance est de plus ou moins 500 Hz. La tranche d'âge 3-10 ans est source de polémique. Les valeurs de fréquence fondamentale varient en fonction des auteurs. Cornut (1991)^[46] estime qu'il n'y a pas de différence significative entre les garçons et les filles et les diverses tranches d'âge. Il rapporte la valeur de plus ou moins 250 Hz. D'autres auteurs découpent la tranche d'âge afin de mieux apprécier les valeurs des fondamentaux par année. Le tableau ci-dessous montrent une grande diversité de valeurs. Pour une même tranche d'âge, une même tâche, nous observons des différences pouvant atteindre 59 Hz entre Ferrand et al (1996)^[80] et Heuillet et al (1997)^[116] pour la tranche 4 ans, ou encore 35 Hz entre Ferrand et al (1996)^[80], Cornut (1991)^[46] et Sorenson (1989)^[292] pour la tranche 7 ans. Par contre, tous s'accordent pour dire que de manière générale le fondamental tend à diminuer avec l'âge.

Tableau 1 : Valeurs de la fréquence fondamentale d'enfants âgés de 3 à 10 ans.

Auteurs	Années	Sujets (N)	Âges	Tâches	Valeurs G	Valeurs F
Ingrisano et al ^[143]	1980	7	4-8	Phrase	323	272
Aronson ^[8]	1985		7		287	
			8		276	
Glaze et al ^[99]	1988	121	5-11	[a]	226	288
Wheat et al ^[344]	1988		6	Spontané	220	211
Sorenson ^[292]	1989	6	6	Spontané	251	296
			7		288	258
			8		229	251
			9		221	266
			10		220	229
Cornut ^[46]	1991	270	3-10	Spontané	250	255
Ferrand et al ^[80]	1996	20	3-4	Spontané	249 (21.4)	257 (10.9)
			5-6		243 (18.9)	256 (27.6)
			7-8		253 (15.7)	234 (37.3)
			9-10		254 (18.3)	240 (15.1)
Heuillet et al ^[116]	1997		4		308	
			5		280	

Par la descente progressive du larynx dans le cou, un abaissement du fondamental usuel s'effectue. Il continuera à décroître discrètement jusqu'à la puberté. Sur le plan anatomophysiologique, cette période se caractérise par une accélération

de la différenciation histologique des cordes vocales. La couche élastique se différencie de la couche de fibres de collagène. Il y a modification de la taille et de la forme du larynx et des cordes vocales. Les cordes vocales gagnent plus ou moins 60% en longueur pour les garçons et 34 % pour les filles (Kahane (1987)^[155]). Le fondamental diminue d'une octave pour les garçons et d'une tierce pour les filles. Ce bouleversement vocal est plus important chez les garçons dont le régime vibratoire des cordes vocales passe du mécanisme léger au mécanisme lourd. Une étude longitudinale d'Hollien et al (1972)^[130] donnent une durée moyenne de 18 mois à la mue avec une étendue de 8 à 26 mois sur une tranche d'âge de 10.5 à 16.5 ans. Sataloff (1997)^[269] mentionne une période de la mue plus active entre 12 ans et demi et 14 ans. La puberté suscite des remises en question tant chez les professeurs que chez le jeune chanteur. Comme le souligne Blatt (1983)^[in 269], la "sagesse" veut que pendant la période de la mue, le chanteur ne travaille pas sa voix. La "science" n'a pas apporté de preuve à cette croyance, sinon que, selon Blatt, continuer à cultiver la voix durant la mue augmenterait les possibilités vocales du sujet. D'autres auteurs comme Andrews (1988)^[in 269], Wilson (1986)^[in 269] ne partagent pas cet avis. Ils mettent en garde contre un entraînement vocal caractérisé par un répertoire trop exigeant, une obligation pour le choriste de conserver son registre aigu, des fluctuations entre grave et aigu, pratiques qui mettraient la voix en danger. Cette mise en garde serait renforcée par une étude exploratoire de Killian (1997)^[163] qui observe les processus de changements vocaux sur une population large de 141 sujets répartis en deux groupes de 77 garçons et de 64 hommes et groupés selon l'entraînement vocal reçu. Interrogés, les chanteurs disent avoir été sensibles à leur mue, et avoir un mauvais souvenir de leur expérience de choriste durant cette période qu'ils qualifient de négatives.

Les valeurs rapportées par la littérature pour les adultes sont fonction de la tâche, du nombre de sujets et du matériel utilisé. Le tableau qui suit reprend quelques valeurs de fondamental.

Tableau 2 : Valeurs de la fréquence fondamentale de sujets adultes âgés de 20 à plus de 65 ans.

Auteurs	Années	Matériel	Sujets	Âges	Tâches	Valeurs H	Valeurs F
Cornut et al ^[46]	1991	GFA ³	45H/45F	20-29	Lecture	119	240
				30-39		119	215
				40-49		121	211
					spontané	108	217
						108	189
						106	188
Morris et al ^[203]	1995	FFI-8 ⁴	50 H	20-35	Lecture	125 (11.1)	
				40-55	[a]	117 (9.4)	
				>65	spontané	130 (16.7)	
Bough et al ^[29]	1996	Fourcin Laryngo ⁵	8H/6F	25-35	[a]15 fois	125 (20.8)	235 (47.4)

Légende : H : hommes ; F : femmes.

Tableau 3 : Moyennes fréquentielles et déviations standards en fonction de trois types d'intensité rapportées par Sulter et al (1996)^[302].

Sulter 1996	INTENSITE		
Hauteur tonale	Faible (SD)	Normale (SD)	Forte (SD)
Hommes : (N=47) 17-35 ans			
Basse	124.3 (17.31)	120.2 (17.40)	126.6 (19.21)
Normale	172.8 (23.43)	171.6 (26.73)	172.3 (26.24)
Haute	249.6 (44.63)	258.7 (56.29)	255.8 (62.26)
Femmes : (N =92) 17-44 ans			
Basse	187.8 (34.46)	203.6 (26.11)	196.0 (27.13)
Normale	259.1 (36.78)	261.1 (36.45)	262.8 (31.83)
Haute	367.8 (61.29)	386.8 (65.41)	376.4 (61.35)

Légende : SD : déviation standard entre parenthèses.

Ferrand et al (1996)^[80] observent qu'en dépit des différentes méthodologies d'analyse et des tâches proposées, les valeurs du fondamental moyen adulte semblent convergentes d'une étude à l'autre. C'est à partir de 50 ans que les valeurs de fréquence fondamentale augmentent graduellement chez l'homme et diminuent chez la femme (Mc Glone et al (1963)^[194]). Les modifications vocales se traduisent en

³ Le GFA.: le "Glottal Frequency Analyser" de Kitzing donne des informations sur la fréquence du son fondamental laryngé (Cornut (1991)^[46]).

⁴ Le FFI 8: le "Fundamental Frequency Indication" est employé pour calculer les fréquences fondamentales parlées (Hollien (1981)^[122]).

⁵ Le laryngographe de Fourcin: utilise les informations recueillies par un électroglottographe. Il permet d'obtenir notamment des histogrammes de la fréquence fondamentale.

général par du souffle sur la voix, une perte de l'étendue, un moins grand contrôle du souffle, une fatigue vocale et une imprécision de la hauteur tonale. Ces manifestations sont associées à des détériorations des fonctions corporelles telles que l'atrophie musculaire et neuronale, l'atrophie ligamentaire, l'ossification des cartilages, la diminution des fonctions respiratoires, certains troubles de l'humeur, une détérioration buccale (perte des dents,...). Le modèle de coalescence (Brown et al (1991)^[32], Hollien (1987)^[135]) décrit ce phénomène. Pour les 7^{ème}, 8^{ème} et 9^{ème} décades, les valeurs de hauteurs tonales se fondent entre les hommes et femmes.

Le terme de presbyphonie⁶ définit le processus de vieillissement de la voix accompagnant le vieillissement général de la personne. Woo et al (1992)^[355] rappellent qu'il ne s'agit pas d'un trouble commun à toutes personnes vieillissantes. Seule une évaluation médicale et vocale peut la diagnostiquer. Jusqu'à présent, aucune valeur de référence n'a pu encore être établie sur "comment" le sujet âgé vieillit. Plusieurs auteurs tentent de définir le vieillissement selon différents concepts. Mishara et al (1984)^[202] tiennent compte de l'âge chronologique, c'est-à-dire du nombre d'années vécues, ils semblent fixer, d'après l'opinion générale et de façon tout arbitraire, le vieillissement à 65 ans. Un fondement aussi discutable peut-il servir de classification?. Meire (1992)^[197] trouve que la vieillesse correspond essentiellement à un critère biologique. C'est le déclin des diverses fonctions vitales qui caractérise le vieillissement de l'organisme vivant, bien que certaines régulations homéostatiques puissent compenser partiellement ces dégradations.

D'autres éléments, comme les facteurs psychiques et sociaux influencent également le vieillissement et se distinguent du vieillissement biologique.

Swine (1992)^[314] décrit le vieillissement comme un processus universel, intrinsèque, progressif, irréversible, et délétère au sens où il s'accompagne de pertes fonctionnelles. Il ajoute que le vieillissement entraîne des modifications qu'il importe de ne pas considérer comme pathologiques. En tant que tel il n'est pas une maladie et ne se soigne donc pas.

⁶ Le terme de **presbyphonie** a été choisi pour caractériser les modifications pathologiques de la voix chez la personne âgée, d'une part parce que ses références étymologiques sont incontestables ; "presbus" signifiant vieux, âgé, et "phonéo", faire entendre un son de voix ; et d'autre part parce que d'autres altérations sensorielles ou fonctionnelles liées à l'âge utilisent la même racine grecque, par exemple, presbytie, presbyacousie, presbyphagie.

La science tente de ralentir ce processus de vieillissement biologique. Les thérapeutes disposent de moyens pour ralentir les effets de l'âge ; les femmes à la ménopause prennent des oestrogènes ; une médication appropriée maintient la fonction musculaire, aide le système vasculaire, nerveux et respiratoire. Pour un professionnel de la voix, des exercices musculaires lui permettent d'entretenir sa condition physique.

Inversement, les chanteurs ne suivent pas ce modèle de coalescence. Les chanteurs de toutes tessitures tendent à maintenir leur niveau moyen de fréquence fondamentale parlée durant toute leur vie (Sataloff (1997)^[269], Hollien (1993)^[136]). Hollien et Shipp, ^[in 69]) comparent leur valeur de fréquence fondamentale parlée en fonction de l'âge et de l'entraînement. Plusieurs études attestent que les locuteurs professionnels plus âgés gardent un modèle vocal étonnamment stable si on le compare à celui de leurs pairs âgés non entraînés (Boone et al (1997)^[27]). Leurs résultats montrent que les groupes professionnels souffrent moins des effets de l'âge que les non professionnels. Leurs conditions vocales telles que l'entraînement quotidien, leur perception fine de la hauteur tonale, leur contrôle de la sensibilité interne les aideraient à maintenir la souplesse de leur musculature laryngée. Ce phénomène présente des pistes thérapeutiques pour le traitement vocal des sujets âgés qui pourraient ainsi contrecarrer les changements vocaux typiques liés à l'âge. Dans l'étude de Morris, les fréquences fondamentales parlées des chanteurs d'âge moyen ne diffèrent pas significativement des chanteurs plus âgés.

Chez les chanteurs âgés, les auteurs observent néanmoins des modifications dans la gestion de l'air, ce qui entraîne de la fatigue vocale et une moins grande précision de la hauteur tonale ; ils notent aussi des changements dans les caractéristiques du vibrato et l'apparition de tremolo (Sindo et al (1990)^[284], Woo et al (1992)^[355], Orlikoff (1990)^[222], Ward et al (1989)^[343], Wilcox et al (1980)^[345], Tanaka et al (1994)^[316]). Timiras (1978)^[319] rapporte une réduction de 40 % de la capacité vocale, s'étalant de 20 à 80 ans.

Ramig insiste sur le fait qu'il faut différencier les déficits vocaux dus à des changements structuraux et physiologiques du phénomène biologique du vieillissement. Des changements vocaux mineurs sont notés chez des individus physiologiquement en forme. Spirodouso (1982)^[295] rapporte que la forme physique

n'est pas seulement un prédicteur de l'intégrité du système nerveux central mais également un marqueur valide de l'âge physiologique. Mais l'âge physiologique peut être plus prédictif que la production vocale de l'âge chronologique (Chodsko et al (1987)^[37]). Il divise la population des sujets âgés de plus de 70 ans en deux groupes selon qu'ils sont en forme ou en méforme. Les changements vocaux normalement attendus pourraient être largement absents chez les sujets en forme et présents chez les sujets en méforme.

Ils émettent l'hypothèse que des locuteurs plus âgés et en forme grâce à l'entraînement et l'expérience vocale peuvent éviter les changements vocaux que subissent ceux qui ne reçoivent pas d'entraînement vocal adéquat.

Morris et al (1995)^[203] comparent un groupe de 50 hommes non-chanteurs à un groupe de 25 chanteurs. Ils mesurent le fondamental moyen et l'intensité sur le "Rainbow passage" [annexe 1], le fondamental moyen en langage spontané et l'étendue fréquentielle sur un [a]. Les résultats montrent que les chanteurs de toutes catégories tendent à maintenir des niveaux de fréquence fondamentale similaires à travers toute leur vie d'adulte. Les chanteurs âgés sont classés à part des non-chanteurs âgés, les hommes et les femmes non-chanteurs indiquant clairement une tendance à changer le fondamental moyen parlé en fonction de leur âge. Ces modifications dues à l'âge ne paraissent pas se produire chez les chanteurs tant sur leur fondamental que sur leur intensité. Les chanteurs de toutes tessitures vocales maintiennent des étendues fréquentielles phonatoires similaires quel que soit leur âge. Ces découvertes contrastent avec la restriction des étendues observée chez les non-chanteurs, hommes et femmes. L'emploi d'une large étendue fréquentielle et le maintien des activités liées au chanteur peuvent contribuer à conserver l'étendue vocale tout au long de la vie.

Morris et al (1995)^[203], quant à eux, formulent trois hypothèses pour expliquer les différences entre chanteurs et non-chanteurs. L'hypothèse de la pression correspond au fait que les chanteurs s'exprimeraient sur une hauteur tonale plus élevée que les non-chanteurs. Cet état de fait serait lié à l'emploi d'un niveau de pression sonore supérieur (Brown et al (1990)^[30], (1991)^[32], (1993)^[33]). L'hypothèse du contrôle traduit le postulat que les chanteurs développent pendant leur pratique un sens interne (somesthésique) de leur fréquence vocale. Ceci leur permettrait de maintenir leur fréquence fondamentale parlée fixe et stable tout au long de leur vie.

L'hypothèse "de l'entretien" serait le fait que la musculature laryngée est entraînée quotidiennement par des exercices isotoniques et isométriques. Ces émissions vocales aideraient à pallier les modifications anatomophysiologiques. [Tableaux 4–8]

Tableau 4 : Valeurs de la fréquence fondamentale pour des chanteurs rapportées par Nadolesny 1923 (in Drew 1995^[69]).

Sujets (N)	Valeurs Femmes			Valeurs Hommes		
	C-A	M-S	S	Basse	Baryt.	Ténor
86F/57H	212	230	262	98	116	134

Légende: C-A: contre-alto, M-S: mezzo-soprano, S: soprano, Baryt.: baryton.

:

Tableau 5 : Valeurs de la fréquence fondamentale pour des non-chanteurs rapportées par Brown et al (1991)^[32].

Sujets	Âges	Valeurs
16	20-35	224
	35-55	218
	65-85	214

Tableau 6 : Valeurs de la fréquence fondamentale pour des chanteurs et des non-chanteurs rapportées par Morris et al (1995)^[203].

Matériel	Sujets (N)		Âges	Valeurs (DS)		
	NCH	CH		NCH	Ténor	Basse
FFI	18	14	20-35	125(11.1)	137(6.9)	112(11.9)
	14	11	40-55	117(9.4)	146(13.6)	127(10.7)
	18	9	>65	130(16.7)	140(8.5)	136(7.5)

Légende: NCH: non-chanteur, CH: chanteur.

Tableau 7 : Valeurs de la fréquence fondamentale pour des chanteuses rapportées par Drew et al (1995)^[69].

Matériel	Sujets	Âges	Tâche	Valeur
CSL (tape)	10 soprani	18-30	Spontané	219

Légende: CSL: Computer Speech Lab de Kay Elemetrics, tape: via un enregistreur.

Tableau 8 : Valeurs de la fréquence fondamentale pour des chanteurs rapportées par Novak et al (1995)^[217].

Matériel	Sujets	Âge	Valeurs femmes			Valeurs hommes		
			A	M-S	S	B	BB	F
3D span (tape)	165	14-17	225	242	255	112	136	168

Légende : 3D span: analyse spectrale en 3 dimensions, (tape): échantillons mesurés via un enregistreur, A: Alto, M-S: mezzo-soprano, S: soprano, B: basse, BB: basse baryton, F: falsetto.

1.2.2.2.1.2. La pathologie.

Murry (1978)^[210] observe que les troubles vocaux organisés ne semblent pas caractérisés par une fréquence fondamentale parlée plus basse que la normale. Cette impression de fréquence fondamentale parlée plus basse est due plus à la perception de la qualité vocale qu'à la seule perception de la hauteur tonale. Galloway et Berry ^[in 210] rapportent que 50 % des sujets qui souffrent d'un trouble vocal ont un problème de hauteur tonale. Sapir et al (1993)^[266] et Cooper (1982)^[45] déclarent que parler trop bas entraîne une dépense d'énergie musculaire laryngée, affecte la voix et engendre de la fatigue et de la raucité. Sapir et al ^[266] remarquent aussi que les mêmes altérations peuvent affecter la voix des chanteurs qui parlent avec un fondamental trop bas. Ils expliquent qu'une interaction optimale entre le larynx et la pression sous-glottique doit exister pour produire une hauteur tonale appropriée au mécanisme donné, et éviter de léser le système. Cette hauteur tonale basse pourrait dès lors être un facteur discriminant pour déterminer le degré de détérioration vocale. Koufman et Blalock (1988)^[168] différencient le syndrome "tension/fatigue laryngée" chez les chanteurs et les locuteurs professionnels. Chez les chanteuses, ce syndrome serait dû aux anomalies de la hauteur tonale et au souffle. Chez les locutrices, il serait dû à des tensions corporelles. Ils spécifient que les valeurs moyennes des fréquences fondamentales parlées des chanteuses sont inférieures à celles des locutrices. Ils incluent dans le syndrome une hauteur tonale inadéquate, un contrôle réduit de la gestion du souffle et une dysphonie. Stemple et al (1995)^[299] font remarquer que la fatigue vocale est une plainte qui accompagne les troubles vocaux. La fatigue pathologique peut être entraînée par la pathologie (œdème, nodules, paralysie) ou la dysfonction. Elle peut être due à de l'abus, à un désordre fonctionnel, à des lésions pathologiques ou neurologiques. Ce trouble s'accompagne d'un effort inhabituel pour

parler et d'une fatigue générale. Greene (1972)^[102] postule que, physiquement, la fatigue vocale correspond à une faiblesse du thyro-aryténoïdien causé par un abus ou un forçage musculaire laryngé. Scherer et Stemple, dans leurs études respectives, provoquent une fatigue vocale chez leurs sujets. Scherer propose une lecture de 15 minutes à intensité forte et l'octave supérieure soit une lecture à 80% de l'intensité maximale tandis que Stemple et al ^[299] proposent deux heures de lecture avec une pause de 5 minutes entre les deux heures. Dans l'étude de Stemple et al ^[299], les 10 femmes âgées de 22 à 45 ans montrent des valeurs moyennes de Fo sur un [a] tenu, significativement plus élevées au post test soit 223 Hz pour 204 Hz au pré test. Ils confortent dès lors l'hypothèse de Greene^[102] concernant l'implication du muscle thyro-aryténoïdien (T-A) dans la fatigue vocale. Lorsque le T-A est sous tension, les sujets éprouvent des difficultés à maintenir une fréquence fondamentale basse. Ils essayent de maintenir leur voix parlée à une hauteur tonale haute, il s'ensuit une fatigue vocale. Sander et al (1983)^[264] lient la fatigue vocale à un déséquilibre entre le soutien respiratoire et l'effort musculaire laryngé. Colton et Casper (1990)^[44] notent que dans les cas de fatigue, l'étendue phonatoire est restreinte, la variabilité de la fréquence fondamentale réduite, la cadence de flux d'air diminuée et la phase de fermeture du cycle vibratoire inadéquate.

Fritzel (in Sataloff ^[269]) observe qu'un des symptômes principaux de l'œdème de cordes vocales est la diminution de la hauteur tonale. Le dysfonctionnement est caractérisé par une réduction de l'efficacité des cordes vocales, une perte des fréquences aiguës, une fatigue vocale et une légère raucité. Dans une étude de Peppard (1988)^[229], les fondamentaux de chanteurs et non-chanteurs porteurs de nodules sont significativement plus bas ($p < 0.005$) que dans le groupe contrôle. La nature des nodules et leurs caractéristiques perceptuelles apparaissent être différentes pour les chanteurs que pour les non-chanteurs. Les locuteurs présentent des nodules de masse plus importante que les chanteurs ($p < 0.001$), une étendue phonatoire moyenne plus basse ($p < 0.001$) et une évaluation anormale de la raucité, non présente chez le chanteur ($p < 0.01$). Les nodules de petite taille du chanteur seraient liés à un fondamental plus bas ; ceux de masse plus importante chez le non-chanteur entraînerait un fondamental plus bas. Drew et al (1995)^[69] mesurent des valeurs de fondamentaux en fonction de la pathologie chez des soprani. Ses résultats

montrent chez elles une fréquence fondamentale parlée plus basse avec fatigue vocale à la fois dans la tâche de lecture et dans le discours spontané.

Tableau 9 : Valeurs de la fréquence fondamentale en fonction de la pathologie.

Auteurs	Années	Sujets	Ages	Tâches	Valeurs F		Valeurs F CH	
					Nl.	Patho	Nl.	Patho.
Peppard [229]	1988	90	18-30	[i]	212	197	244	207
Drew ^[69]	1995	18		Lecture			230.4	201.8
				Spontané			218.6	109.9

Légende : Nl.: normale, Patho: pathologique.

1.2.2.2.1.3. La culture.

Lorsque que Drew^[69] compare les valeurs moyennes des fondamentaux de ses sujets avec celles de Nadolesny, il observe, chez les femmes, des valeurs plus basses (218Hz versus 262Hz). Une différence culturelle pourrait-elle expliquer ce fait, au moins en partie ?. Luchsinger et Arnold (1965)^[187] suggèrent que comparées à des femmes d'autres cultures, les Américaines trouvent désirable de s'exprimer sur une hauteur tonale basse. Koufman et Blalock (1988)^[168] adoptent cette hypothèse. Les Américaines estiment qu'une hauteur tonale basse rend la voix plus dramatique, sensuelle et moderne. Ils baptisent ce fait le "Bogart-Bacall syndrom".

Sapienza (1997)^[265] remarque que la plupart des valeurs de référence sont établies sur des populations blanches. Elle estime que si des distinctions fonctionnelles existent entre des groupes raciaux, il faut en tenir compte dans les modèles courants qui caractérisent la production vocale normale et les intégrer aux bases de données. Jusqu'à présent, aucune différence significative n'est rapportée. Sapienza rapporte une observation anatomophysiologique de Boshoff portant sur 102 larynx de noirs africains et 23 larynx de caucasiens. Elle conclue que les noirs africains disposent d'une musculature laryngée plus importante. Ces différences anatomiques renforcent l'auteur dans l'idée qu'il pourrait y avoir des diversités de fonctionnement vocal. Son groupe sujet est constitué de 20 Américains blancs et de 20 Afro-américains répartis selon le sexe, la grandeur et le poids, âgés de 10 à 28

ans. Aucune différence n'est observée. Les fondamentaux sont respectivement 123.85 Hz pour les Afro-américains et 124.53 Hz pour les blancs, et 223.10 Hz pour les Afro-américaines et 214.90 pour les blanches. Elle reconnaît que les deux groupes, tous deux baignés dans le milieu américain, se différencient plus par leurs origines que par leur culture.

D'après Mayo et al (1994)^[191] lors d'un jugement perceptuel d'identification de sujets blancs et noirs, les juges sont capables de différencier les deux groupes raciaux avec une haute précision à partir de la tenue de courte durée d'une simple voyelle. En 1995^[192], ils émettent l'hypothèse que cette reconnaissance serait liée aux formants fréquentiels de la voyelle différents entre les deux races. L'hypothèse est vérifiée sur 40 sujets afro-américains et blancs. La fréquence des 1^{er}, 2^{ème} et 3^{ème} formants est significativement plus basse pour les locuteurs afro-américains. Ils attribuent ces différences acoustiques aux caractéristiques céphalo-faciales. Mais Walton et Orlikoff (1994)^[340] ne trouvent pas de différences formantiques chez 50 locuteurs Afro-américains et blancs. Roberts (1966)^[257], Lass (1979)^[176] restent convaincus que la fréquence fondamentale serait un facteur déterminant dans l'identification du locuteur. Cependant jusqu'à présent, aucune donnée ne permet de le confirmer.

1.2.2.2.1.4. L'état émotionnel.

Les relations entre la voix et l'état émotionnel sont de nature complexe. Les ouvrages regorgent de réflexions et de propos à ce sujet. Les auteurs éprouvent des difficultés à réaliser des études scientifiques. Nilsonne, in Drew et al^[69], trouve pour 16 patients dépressifs une corrélation entre les mesures du fondamental (déviations standard, cadence des changements de la fréquence fondamentale) et l'état clinique des patients. Par contre la fréquence fondamentale ne diffère pas significativement.

Scherer (1987)^[274] suggère que les signaux vocaux sont d'importants indicateurs des états émotionnels et en particulier du stress. L'état émotionnel joue un rôle sur la respiration et la tension musculaire et donc fatalement sur le processus vocal. Son implication dans la phonation et l'articulation modifierait les paramètres acoustiques de la voix. L'analyse acoustique chez des sujets stressés révèle des différences de fréquences fondamentales, (Ekman et al (1976)^[72]) attribuées au

mécanisme de "coping"⁷, lequel diffère d'un individu à l'autre (Scherer (1981)^[273]). En conséquence, il est permis de supposer diverses réponses psychologiques et comportementales entre individus dans la gestion du stress. Deux tendances en terme de stratégies de défense apparaissent, l'une peut être fonctionnelle et gérer le stress, l'autre peut être dysfonctionnelle et exacerber le stress. La variabilité de la fréquence fondamentale jouerait un rôle essentiel dans la corrélation entre la stratégie de "coping" découlant de la personnalité et la mesure de la qualité vocale. Mc Hugh et al (1997)^[196] rapportent des valeurs de fondamentaux très variables chez des sujets dont le mécanisme de défense est émotionnel. Ils s'expriment avec plus d'inflexions dans la voix, moins de régularité dans leur phonation. Cependant, la valeur de leur Fo n'est pas significativement différente. Déjà en 1978, Murry^[210] montrait des résultats similaires.

Il apparaît donc que l'émotion peut modifier momentanément la mélodie et la hauteur de la voix ; cependant elle ne change pas la hauteur fondamentale.

L'observation des traits psychologiques de sujets dysphoniques, porteurs de nodules indique des corrélations significatives entre les états émotionnels et la pathologie sont significatives. Heaver (1958)^[in 114] trouve un certain nombre de facteurs psycho-pathologiques chez les patients dysphoniques. Les nodules vocaux peuvent être diagnostiqués chez des adultes dont certains traits psychologiques génèrent un état de tension, de la colère ou de la dépression. Ces traits comprennent la logorrhée, un état tendu, des interactions sociales agressives, des problèmes interpersonnels aigus ou chroniques (Aronson (1990)^[9]). Chez les enfants, des études psychiatriques (Wilson (1973)^[346], Nemec (1961)^[213]) révèlent également des traits de personnalité caractéristiques chez les sujets porteurs de nodules. Ceux-ci sont décrits comme manifestant de la frustration, des tensions ou encore des troubles émotionnels. Greene (1989) étudie 60 enfants répartis en deux groupes soit 30 sujets sains et 30 sujets porteurs de nodules. Les résultats l'amènent à conclure que certains traits psycho-comportementaux caractérisent les enfants avec nodules. Plus de 60% des enfants sont évalués dans les termes suivants : "hystériques", "ayant toujours le dernier mot", "désobéissants voire menaçants", "se plaignant d'injustice", "troublés

⁷ Le "coping" selon Lazarus (1990)^[179] représente "les efforts de modifications cognitives et comportementales constantes face à des demandes externes et internes évaluées comme stressantes ou dépassant les ressources du sujet".

dans leur relation avec leurs pairs". Cette étude renforce le lien entre les caractéristiques psycho-comportementales et la voix. Morrison et al (1983)^[204] suggèrent un lien entre les stratégies de "coping" et la dysphonie. Ils estiment que les patients souffrant d'une dysphonie due à des tensions musculaires présentent des difficultés à gérer leur stress. Egger et al (1990)^[70], Freidl et al (1989)^[84] corroborent ces constats. Ils rapportent que leurs sujets dysphoniques montrent des niveaux plus élevés d'anxiété face à des situations stressantes. Ils soulignent que les événements ont une influence importante dans l'étiopathogénèse de la dysphonie. Pour Mans (1992)^[188], le stress diagnostiqué au début de la dysphonie est d'une importance primordiale dans le développement de celle-ci également pour un groupe de patients souffrant de granulome.

En conclusion, il y a une corrélation évidente entre le caractère du patient et la prépondérance à développer des lésions consécutives au malmenage comme les nodules ou le granulome. Cependant, ces études sont à nuancer. La majorité reposent sur des bases empiriques et restent donc sujettes à caution.

1.2.2.2.1.5. Les facteurs hormonaux.

L'hypothyroïdisme est reconnu comme entraînant des difficultés vocales telles que raucité, fatigue vocale, voix étouffée, perte d'étendue, sensation de voix rocailleuse (laryngite myxoedémateuse). Néanmoins, le mécanisme des modifications vocales associées à l'hypo-thyroïdisme n'est pas encore complètement compris. Pourtant, il y 30 ans, Ritter^[in 269] démontrait, chez ces patients, un niveau d'acide mucopolysaccharide sous-muqueux présent dans la corde vocale. Cet acide agit sans doute comme un diurétique osmotique et augmente le contenu fluide dans la lamina propria. La masse de la corde vocale est alors augmentée et la vibration cordale diminuée. Cela peut s'apparenter à l'œdème de Reinke. Un hypo-thyroïdisme sévère avec myxoedème entraîne des changements plus profonds associés parfois à une chute de la force musculaire et une paralysie des cordes vocales.

La prise de stéroïdes anabolisants provoque des modifications vocales (Vuorenkoski (1978)^[338]). Chez les hommes, cette médication entraîne un abaissement de la fréquence fondamentale et parfois une rugosité dans la voix. Quant

aux voix féminines, elles se caractérisent par une virilisation vocale (fréquence fondamentale nettement diminuée jusqu'à une octave) dont l'effet est irréversible.

Après **la ménopause**, la fréquence fondamentale chute. Les ovaires continuent à sécréter des androgènes mais plus d'œstrogènes. Les facteurs affectant la voix correspondent à une diminution de la puissance pulmonaire, une atrophie des muscles laryngés, un amincissement des cordes vocales, une perte des fibres élastiques et collagènes. On assiste à une virilisation de la voix ; la voix est soufflée. L'étendue globale vocale est réduite, et particulièrement dans les fréquences aiguës. Les chanteuses peuvent développer un trémolo et une imprécision dans la hauteur tonale.

Les variations vocales lors du **cycle menstruel** sont difficiles à quantifier. La plupart des effets néfastes sont observés au 25^{ème} jour du cycle. Le trouble vocal est caractérisé par une décroissance de l'efficacité vocale, une perte des fréquences aiguës, une fatigue vocale, une légère raucité, une voix plus sombre. Ces symptômes sont plus apparents chez les chanteuses que chez les locutrices. Ces mêmes symptômes sont observés au 14^{ème} jour du cycle pendant l'ovulation. (Higgins et al (1991)^[120]). Selon Abitbol et al (1999)^[2], 33% des femmes souffrent de ce syndrome. La cause de ces dysphonies n'est pas encore complètement comprise. L'activité prémenstruelle combinée progestérone/œstrogène provoque une vasodilatation, d'où un engorgement des vaisseaux et un œdème de la corde vocale. Isenberg ^[in 269] a recherché les liens possibles entre le système reproducteur et la voix chantée. Il étudie le cas des chanteuses souffrant d'aménorrhée ; elles se plaignent de souffle sur la voix, de faiblesse vocale, d'une incapacité à produire certaines hauteurs tonales, à assouplir le passage de la voix de poitrine à la voix de tête, d'une difficulté à réaliser des traits musicaux rapides. Ces plaintes sont similaires à celles des chanteuses durant leur période prémenstruelle. Abitbol étudie l'effet d'une thérapie médicamenteuse sur 97 femmes en période prémenstruelle âgées de 23 à 36 ans ; chez 84 d'entre-elles, il a constaté une amélioration.

Les **femmes enceintes** peuvent également souffrir de modifications vocales du même ordre que celles subies lors des troubles prémenstruels. Les changements peuvent en partie s'associer à l'œdème de la couche superficielle de la lamina propria

provoqué par une augmentation importante des niveaux d'œstrogène et de progestérone.

1.2.2.2.1.6. Le sexe.

Peu d'études se sont penchées sur les différences sexuelles entre les voix d'enfants, et celles d'adultes en se basant sur un large éventail de paramètres acoustiques. La fréquence fondamentale moyenne est la mesure la plus courante (Ferrand et al 1996 ^[80]). Il est d'ailleurs généralement admis que la fréquence fondamentale parlée soit le marqueur le plus révélateur de la perception de la voix (Avery et al (1993)^[11]). Les études qui suivent nuancent cependant cette conception. Andrews et al (1997)^[6] tentent d'identifier les caractéristiques perceptuelles et acoustiques d'un groupe de 11 hommes hétérosexuels travestis âgés de 35 à 70 ans et ne prenant pas d'hormones. Quatre-vingt-huit juges ont évalué les voix à l'aide d'une échelle perceptuelle composée de 12 adjectifs. Les sujets devaient faire une lecture sur deux modes, l'un avec leur voix masculine, l'autre avec leur voix féminine. Chacune des voix a été significativement perçue plus féminine lorsque le locuteur optait pour le mode féminin. La hauteur tonale, le souffle ⁸, la "mélodie" vocale (son caractère animé) se sont avérés des paramètres particulièrement liés à la perception du caractère féminin de la voix. Le paramètre "soupiré" sonnait plus clair, est aussi un des plus imités par le locuteur masculin. Le caractère "animé" suggère une plus grande variabilité vocale du locuteur. Les paramètres acoustiques de cette étude, ils ne vérifient pas l'hypothèse que la variabilité entre les deux modes soit purement fondée sur la fréquence fondamentale parlée. Ce sont plutôt les variations obtenues dans les paramètres de puissance, cadence, durée et contour "intonatoire" qui seraient plus accentuées. Les locuteurs ralentissent leur cadence, allègent l'intensité et modifient leur contour "intonatoire" comme un exercice stylistique. Il est intéressant de noter que ce sont les deux sujets dont les fréquences fondamentales étaient les plus basses soit 106 et 105 Hz qui produisaient les plus grandes différences entre les deux modes, et ce, particulièrement pour les paramètres de cadence et de durée. Les sujets chez qui le niveau de hauteur tonale était plus important montraient un mode

⁸ Le souffle n'est pas considéré ici au sens pathologique du terme mais comme une voix "aérienne", "soupirée".

féminin moins convaincant pour les juges. Ceux-ci notaient un manque de flexibilité des caractéristiques de durée, de cadence et d'intonation. Les résultats d'Andrews et al^[6] corroborent ceux de Wolfe et al (1990)^[351]. Dans son étude, elle enregistre des questions-réponses chez 20 sujets transsexuels. Un jury composé de 10 étudiants évalue le sexe de chaque locuteur. Un autre composé de 8 professionnels évalue chaque échantillon sur une échelle de féminité graduée de 1 à 7. L'auteur constate que les contours "intonatoires" ascendants sont perçus comme plus féminins. Une des locutrices transsexuelles est perceptuellement évaluée comme femme malgré une fréquence fondamentale parlée très basse. Les modèles "intonatoires" de cette locutrice sont caractérisés par un pourcentage élevé d'inflexions ascendantes et descendantes. Deux autres de ces locutrices ayant des niveaux similaires de fréquence fondamentale parlée mais moins d'inflexions "intonatoires" ont obtenu des scores plus bas sur l'échelle de la féminité. Néanmoins, une différence majeure est observée entre ces deux auteurs, les juges de l'étude d'Andrews^[6] portent leur évaluation sur la féminité ou la masculinité du sujet. Par contre, dans l'étude de Wolfe, les juges ont évalué l'échantillon comme voix de femme ou d'homme. Gelfer et al (2000)^[90] font remarquer la possibilité que les juges aient perçu tous les sujets de l'étude d'Andrews comme des hommes dont la voix semblait ou plus féminine ou plus masculine.

Avery et al (1993)^[11] concluent également que l'intonation est un marqueur suprasegmental de la parole qui diffère selon les modèles féminins et masculins.

Plus récemment, Gelfer et al (2000)^[90] analyse la fréquence fondamentale parlée, la limite supérieure et inférieure de la fréquence fondamentale parlée, les modèles intonatoires, et les fréquences des formants des voyelles chez 15 sujets transsexuels devenus femmes. Six femmes et trois hommes constituent le groupe de contrôle. Les sujets devaient lire et émettre les voyelles [a] et [i]. Vingt juges évaluent le degré du caractère féminin ou masculin de la voix, caractère que l'auteur corrèle avec paramètres acoustiques. Gelfer et al rapportent des différences significatives ($p < 0.05$) pour la fréquence fondamentale parlée et la limite supérieure de cette fréquence. Les sujets perçus comme "femmes" ont une fréquence fondamentale parlée et une limite supérieure plus élevées que les sujets perçus comme "hommes". Les auteurs trouvent également une corrélation entre les valeurs de la limite supérieure de la fréquence fondamentale parlée et le degré de féminité/masculinité.

Plus la limite est élevée, au plus le juge évalue le sujet avec un degré élevé de féminité. Les auteurs soulignent, néanmoins, la présence d'un biais : la disparité entre les sujets transsexuels. Certains sont des transsexuelles accomplies ; d'autres sont au début de leur vie de femme. Certains ont bénéficié d'une thérapie vocale, d'autres pas. L'âge des sujets variaient de 22 à 63 ans. Dès lors, les juges ont peut-être employé des critères différents pour déterminer le sexe des sujets qu'ils perçoivent comme étant jeunes ou âgés. Un contrôle plus rigoureux de ces facteurs clarifierait ces résultats. Aucun sujet perçu comme femme ne présente une fréquence fondamentale parlée en-dessous de 156 Hz.

Wolfe et al (1990)^[351] corroborent le résultat de Spencer (1988)^[294] qui précédemment constatait que les locutrices dont le fondamental était supérieur à 156-160 Hz étaient perçues comme femmes. Elle ajoute que s'il est nécessaire pour un sujet transsexuel d'atteindre une fréquence fondamentale parlée supérieure à 156 Hz, cette caractéristique n'est cependant pas suffisante : les contrastes "intonatoires" sont essentiels.

Tableau 10 : Valeurs de la fréquence fondamentale parlée selon différents auteurs.

Auteurs	Années	Sujets	Échelle perc*.	Valeurs Fo	
				Pré-op**	Post-op***
Bralley ^[in 351]	1978	Transsexuel		145	165
		Échelle	1 à 7	2.0-5.9	3.7-4.6
Mount ^[in 351]	1988	Transsexuel		110	210
Gelfer et al ^[90]	2000	15 transsex.			164-199
Spencer et al ^[294]	1988	8 transsex.			165-209
				Perçu H	Perçu F
Wolfe et al ^[351]	1990	20 transsex.		118	156-195
Andrews et al ^[6]	1997	11 travestis		119	135

Légende : *perc. : perceptuelle ; **pré- op : pré- opératoire ; ***post- op : post- opératoire.

1.2.2.2.1.7. La posture.

Comme nous l'avons mentionné au paragraphe 1.1.1., l'attitude et la posture jouent un rôle important dans le fonctionnement vocal. Les muscles crico-thyroïdiens tendent les cordes vocales. Le degré d'activité de ces muscles est corrélé significativement à l'augmentation de la hauteur tonale de la voyelle. Lorsque le larynx s'élève, la fréquence fondamentale augmente. Cette élévation du larynx est associée

à l'activité des muscles sus-hyoïdiens (géniohyoïde, du génioglosse et des muscles périlaryngés). Elle entraîne donc à son tour l'os hyoïde et le cartilage thyroïde bascule vers l'avant. Cette inclinaison augmente la longueur des cordes vocales et les met sous tension, provoquant une élévation du Fo. Le degré d'aperture buccale a un impact sur le cartilage thyroïde et joue un rôle non négligeable dans le mouvement de bascule.

Lin et al (2000)^[184] observent l'effet de l'extension de la tête et de l'avancée linguale sur les mesures du Fo et de leur perturbation. Ils rapportent des valeurs de Fo augmentées chez l'homme et la femme suite à un mouvement d'extension de la tête. Les valeurs de Fo sont, par contre, diminuées lors de l'étirement de la langue. Si le mouvement d'extension de la tête s'ajoute à l'avancée linguale, les valeurs de fréquence fondamentale sont diminuées pour les hommes et presque identiques pour les femmes. L'extension de la tête a néanmoins un effet plus important sur la Fo que l'avancée linguale. Ceci suppose que la position complexe hyoïde/larynx est plus susceptible d'être affectée par l'extension de la tête.

Tableau 11 : Valeurs de la fréquence fondamentale en fonction de la posture de la tête et des mouvements linguaux rapportées par Lin et al 2000^[229].

Positions	Valeurs Hommes (N=66)	Valeurs Femmes (N=49)
Tête nle/ langue nle	122.13 (22.39)	205.97 (29.82)
Tête nle/ langue avancée	119.96 (21.57)	203.26 (34.31)
Tête extension/ langue nle	126.43 (24.40)	209.50 (30.26)
Tête extension/ langue av.	124.98 (22.76)	209.10 (30.50)

Légende : nle : position de la langue au repos, AV. : avancée.

En voix chantée, si la bouche s'ouvre au fur et à mesure que le sujet monte la gamme, c'est en réponse au besoin d'harmoniser les fréquences et d'uniformiser le timbre, lorsque la fréquence fondamentale chantée est plus haute que le 1^{er} formant de la plupart des voyelles. Sundberg ^[in 269] explique ce procédé : en déplaçant vers l'aigu la fréquence du premier formant (en ouvrant la bouche), le chanteur permet au formant de venir accroître l'amplitude du son de Fo. Ceci dit, l'ouverture de la

mâchoire n'est pas le seul moyen pour renforcer la fréquence du 1^{er} formant. Le raccourcissement du conduit vocal obtenu en tirant les commissures des lèvres en arrière provoque un effet similaire.

1.2.2.3. L'intensité vocale

1.2.2.3.1. *Le Mécanisme de contrôle*

L'intensité vocale est fonction du degré d'excitation transmise par les ondes glottiques à l'air du tractus vocal. Toutes choses étant égales, plus grande est l'amplitude de l'onde glottique, plus grande est l'excitation qui en résulte au niveau du tractus vocal et plus intense est le signal vocal.

En augmentant la pression de l'air expiré, on accroît la quantité d'air qui est poussée à travers la glotte chaque fois qu'elle est ouverte, et on obtient ainsi une onde glottique plus grande. Plus la pression pulmonaire augmente et plus l'intensité s'accroît. Toutefois, nous savons que c'est l'arrêt brutal du flux qui est la cause première de la mise en vibration acoustique de l'air dans le tractus vocal. Aussi, plus cet arrêt est net et rapide, plus l'excitation dans le tractus est grande et plus le signal est intense. L'intensité n'est pas seulement obtenue par une augmentation de la pression, mais aussi par un ajustement volontaire de la biomécanique des cordes vocales pour renforcer la résistance du flux d'air que le degré accru de pression va produire. Ici, l'augmentation d'intensité résulte d'une interaction très précise entre une augmentation de pression et une augmentation de la résistance glottique au flux d'air.

1.2.2.3.2. *La perception de l'intensité*

La pression acoustique s'exprime par le logarithme du rapport entre la pression acoustique du son et la pression acoustique de référence.

$$\text{SPL} = 20 \log_{10} \frac{P}{P_0} \text{ dB}$$

Sur le plan du spectre vocal, le niveau sonore est défini par l'équation :

$$P_{\text{tot}} = (P_1^2 + \dots + P_n^2)^{1/2}$$

P_{tot} correspond à la pression sonore totale et $P_1, P_2, P_3, P_n \dots$ sont les niveaux de pression sonore des partiels ou harmoniques. Il existe deux systèmes pour nommer les composantes de fréquences sinusoïdales d'un son périodique. Le numéro de l'harmonique ou du partiel est le même que celui de la fréquence relative soit $5F_0$ est le 5^{ème} harmonique ou le 5^{ème} partiel. Pierce (1983)^[233] emploie le terme d'harmoniques pour des sons strictement purs.

Si on considère que l'amplitude du partiel le plus fort du spectre est à 1.0, l'amplitude des autres partiels est inférieure à 1.0. Leur valeur au carré est donc faible et cela d'autant que leur valeur est inférieure à 1.0. En fait, le niveau de pression sonore d'un spectre est pratiquement déterminé par l'amplitude des 2 ou 3 partiels les plus forts. Gramming et al (1988)^[101], Sundberg (1994)^[311] et Titze (1992)^[324] ont d'ailleurs démontré que le niveau de la pression sonore lors de l'émission d'une voyelle est déterminé par le partiel le plus fort, qui est considéré comme la fréquence fondamentale. C'est vrai pour les hauteurs tonales élevées telles que celles que chante une soprano. C'est également vrai lors d'une phonation de faible intensité lorsque les amplitudes sont faibles et que le partiel au sommet du premier formant est plus faible que la fréquence fondamentale. Habituellement le partiel le plus ample dans le spectre d'une voyelle participe au 1^{er} formant. Les amplitudes des partiels du spectre de la voyelle sont déterminées par 2 facteurs. Le premier est la répartition des partiels dans le spectre, lesquels sont contrôlés par l'articulation et la longueur du tractus vocal (Sundberg (1991)^[308]). Le second facteur est le spectre de la source vocale. Les amplitudes des "overtones" ou partiels plus amples que la fréquence fondamentale sont déterminées par les arrêts brusques de l'onde glottique. A son tour, celle-ci dépend de la pression sous-glottique et de la résistance glottique. A l'émission, si le son n'est pas interrompu ou réfléchi, son intensité perd 6 dB à chaque doublement de la distance.

Dans la plupart des cas, il y a correspondance entre la puissance vocale perçue par l'auditeur et la pression acoustique. La puissance est exprimée également en dB. Spontanément on considère que le chanteur a une puissance supérieure au non-chanteur. Pourtant leur pression sonore n'est pas significativement différente. La différence perçue provient de la capacité à produire ces partiels supérieurs à la fréquence fondamentale ou "overtones".

1.2.2.3.3. L'intensité en fonction de la fréquence

Le locuteur ajuste son niveau d'intensité en fonction de l'environnement sonore, de la distance des auditeurs, et du nombre d'auditeurs. Un locuteur sans entraînement vocal éprouve des difficultés à dissocier l'augmentation de la pression sous-glottique et la résistance glottique, aussi la fréquence tend-elle à s'élever parallèlement à l'intensité. Les valeurs de la fréquence fondamentale et de l'intensité sont corrélées. Le chanteur entraîné dissocie parfaitement la gestion de la résistance glottique et l'augmentation de la pression sous-glottique (Boone et al (1994)^[26]).

1.2.2.3.4. L'effet Lombard

L'effet Lombard consiste pour le sujet à développer une intensité plus forte lorsqu'il se trouve dans un environnement sonore qui couvre sa propre émission. Dans une chorale, il peut arriver qu'une voix soit masquée par les voix environnantes. Dans ce cas, le choriste pallie ce manque de feed-back auditif en forçant sa voix. Foot (1965)^[82] a observé que de nombreuses chorales avaient tendance à chanter à des niveaux sonores élevés.

Sataloff ^[269] a émis l'hypothèse que cette compensation serait plus fréquente chez les choristes que chez les solistes qualifiés, qui développeraient un meilleur contrôle de leur voix. Les causes de l'effet Lombard ne sont pas encore complètement comprises. Lane et al (1971)^[174] postulent qu'il s'agit pour le locuteur d'être compris par l'interlocuteur. D'autres imaginent qu'il s'agit plutôt d'un auto-contrôle. Coleman et al (1978)^[42] ont étudié plus spécifiquement ce phénomène. Dans leur étude, les sujets

devaient chanter avec accompagnement via un casque. Ce sont les choristes qui décidaient du niveau d'intensité sonore adéquat. Les accompagnements enregistrés étaient diffusés à, respectivement 90, 95 et 110 dB. Les résultats ont montré que les chanteurs inexpérimentés ne variaient pas autant leur intensité vocale que les chanteurs expérimentés. Ekstrom (1960)^[73] et Maxwell (1975)^[190] supposent que les chanteurs inexpérimentés dépendent davantage de leur feed-back auditif. Ces résultats contradictoires s'expliqueraient par une plus grande maîtrise de la part des chanteurs les mieux entraînés. La directive "chanter à un niveau sonore adéquat" les aurait encouragés à maintenir une intensité légèrement supérieure. Durant la phonation, il existe également un feed-back tactile de la gorge et des muscles concernés qui jouerait un rôle dans l'adéquation du rendement vocal. Scotto di Carlo et al (1994)^[281] parlent de sensibilité de nature palloesthésique (vibratoire) et kinesthésique (musculaire). Ces sensibilités fournissent aux chanteurs des points de repère précis et des moyens de contrôle de leur émission plus fiables que le feed-back auditif qui subit d'importantes modifications en fonction de l'acoustique de la salle. Le chanteur développe cette stratégie pour parer au "bruit" de l'orchestre ou d'un groupe de choristes. Tant les chanteurs que les orateurs utilisent et cultivent ce feed-back tactile de manière efficace. Cependant, le mécanisme et le rôle de ce feed-back tactile n'est pas complètement compris. Les scientifiques ont longtemps ignoré cette sensibilité interne. Seul Husson (1960)^[142] a examiné la sensibilité interne dans la zone palatale. Il a stimulé la zone électriquement ou au contraire l'a anesthésiée en la refroidissant au chlorure d'éthyle. Il en a résulté un manque d'ajustement sur le plan glottique se manifestant chez le chanteur par une altération de la qualité vocale.

Elliott et al (1995)^[75] tentent d'éclaircir la controverse. Ils comparent différents niveaux d'intensité vocale chez 27 chanteurs âgés de 20 à 64 ans en fonction de leur degré d'entraînement soit de 0 à 41 ans de pratique vocale. Les chanteurs sont répartis en trois groupes. Le 1^{er} groupe comprend les sujets très entraînés, le 2^{ème} les sujets moyennement entraînés et le 3^{ème} les sujets peu entraînés ou groupe contrôle. Les chanteurs reçoivent pour instruction de résister à l'effet Lombard. Les mesures d'intensité sont calculées avant et après instruction. Il constate que la plupart des chanteurs quelle que soit leur expérience ont tous tendance à céder à l'effet Lombard lorsqu'ils perdent leur feed-back auditif. Ils sont par contre capables de contrôler leur niveau d'intensité vocale après quelques instructions (85dB en pré- test versus 82dB en post- test ; $p < 0.001$).

1.2.2.3.5. Le niveau d'intensité.

L'emploi d'une intensité excessive peut contribuer à l'apparition de nodules, de polypes, de laryngites chroniques, ou encore d'autres pathologies (Case (1991)^[36], Prater et al (1984)^[237] et Colton et al (1990)^[44]).

Pour Gelfer et al (1997)^[89], les locuteurs sains usent d'une étendue d'intensité moins stressante pour le mécanisme laryngé et moins éprouvante pour le système respiratoire. Les auteurs confirment d'ailleurs que les abus vocaux seraient dus à l'emploi de niveaux d'intensité plus élevés, cause de plus de stress sur le mécanisme phonatoire. Ils fournissent des valeurs de référence sur les caractéristiques de l'intensité de 40 jeunes adultes.

L'intensité dépend également du sexe et dans son étude, les mesures d'intensité des hommes dépassent significativement celles des femmes de 2 dB. La limite supérieure de l'étendue d'intensité pour les hommes est de 79.96 dB et pour les femmes 77.23 dB. Les deux groupes accentuent le caractère fort ou faible de l'intensité au re-test ; ainsi si la tâche est à intensité faible, ils la donneront plus faible (59.25 DB vs 60.79, $p=0.001$) et dans l'autre cas plus forte.

Coleman et al (1977)^[41] et Brown et al (1993)^[33] avaient déjà étudié la fiabilité de l'intensité dans le temps. Coleman et al demandaient au sujet de répéter la voyelle [a] et de courtes phrases à 48 heures d'intervalle. La différence entre les deux productions était de 2dB. Brown et al ,eux, faisaient répéter la tâche 5 jours de suite. La moyenne des variations était de 6.4 à 8.7 dB avec une plus grande variabilité dans la répétition d'une phrase.

Tableau 12 : Valeurs d'intensité selon différents auteurs.

Auteurs	Années	Sujets	Âges	Niveaux Intensité à 30 cm*
Colcord et al ^[40]	1979	1F/ 7H	23.2	70.4 dB
Brown et al ^[33]	1993	14 chanteurs	20-55	61.9 dB
		10 non-chanteurs	20-55	64.7 dB
Morris et al ^[203]	1995	14 chanteurs	20-35	67 dB
		18 non-chanteurs	20-35	71.6 dB
Gelfer et al ^[89]	1997	20 F	20-30	68.15 dB
		20 H	20-30	70.42 dB

Légende: H: hommes; F: femmes; *:niveaux d'intensité pris à 30 cm de la bouche du sujet.

1.2.2.4. Le timbre

Le son produit par les cordes vocales en vibration est un son complexe. Il contient une fréquence de base, le fondamental, et des harmoniques. L'amplitude de ces harmoniques diminuent de plus ou moins 12dB par octave (Pierce (1983)^[233]). Cependant des différences de qualité vocale sont présentes à mesure que le son produit à la source passe à travers le conduit vocal supra-glottique. Le pharynx et la cavité orale connectés à la cavité nasale constituent un système de résonance complexe. La forme de ces structures peut changer et les résonateurs varier. Certaines fréquences dans le conduit vocal sont atténuées et d'autres renforcées. Le conduit vocal comporte plusieurs résonances qui font ressortir diverses bandes de fréquences dans le son produit par la vibration des cordes vocales. Les résonances ont donc un effet sur le timbre des sons musicaux. Une modification de la forme du conduit vocal, entraîne une modification des fréquences de résonances qui déterminent le son produit.

Les fréquences augmentées sont diffusées avec des amplitudes proportionnellement plus élevées. Fant (1973)^[79] et Sundberg (1970)^[303] ont démontré que le conduit vocal comporte 4 ou 5 fréquences de résonance très importantes nommées formant et dont la présence modifie le spectre vocal [Figures 9 et 10].

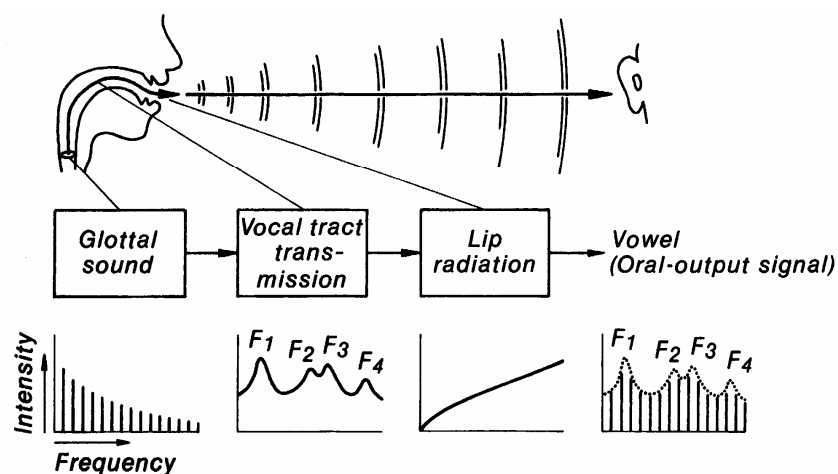


Figure 9 : Facteurs déterminants le spectre de la voyelle.

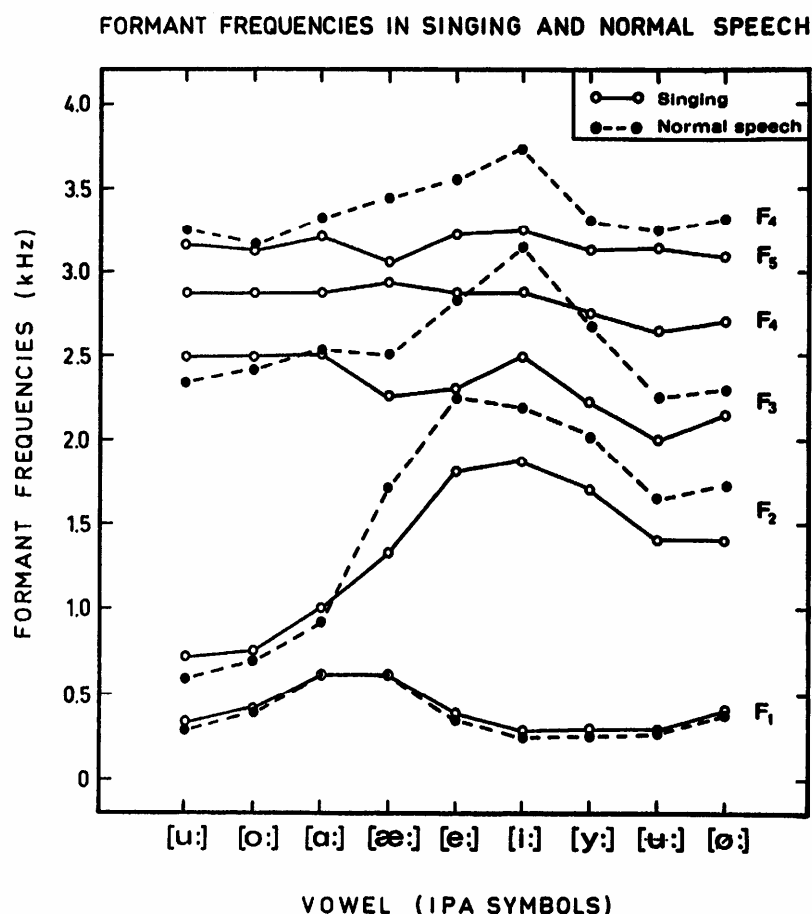


Figure 10 : Formants des voyelles en voix parlée et chantée.

Ces modifications de l'enveloppe spectrale sont responsables de la sonorité parlée ou chantée. Les fréquences des formants sont établies par la forme du conduit vocal, elle-même altérée par la musculature laryngée, pharyngée et la cavité buccale, bien que la longueur et la forme du tractus vocal soient individuellement fixées par l'âge et le sexe de l'individu (les femmes et les enfants disposent de conduits vocaux plus courts et d'un nombre de formants plus élevé). Une réduction de l'ouverture labiale et un allongement du tractus vocal en abaissant le larynx ou en allongeant les lèvres abaissent tous les formants fréquentiels. A l'inverse, la constriction du tractus vocal dans la région glottique entraîne une augmentation des formants fréquentiels. La mandibule peut agrandir le tractus vocal dans la région labiale ou le contracter dans la zone laryngée, la fréquence du 1^{er} formant en est augmentée. Le 2^{ème} formant est particulièrement sensible à la forme de la langue ; le 3^{ème} formant à la position haute de la langue et à la place qu'elle occupe dans la bouche. La relation entre la forme du

tractus vocal et les 4^{ème} et 5^{ème} formants est plus complexe et difficile à contrôler. Elle semble néanmoins dépendre de la longueur du tractus et de la configuration de la profondeur du pharynx.

La maîtrise dans l'ajustement de la forme du tractus vocal est fondamentale dans un programme d'entraînement vocal. Bien que les formants diffèrent pour chaque voyelle, une zone de fréquence a retenu l'attention des chercheurs [63, 78, 112, 198, 308, 325] : le formant du chanteur. Ce formant apparaît aux alentours de 2300 à 3200 Hz. Les 3^{ème}, 4^{ème} et 5^{ème} formants, proches en fréquence forment un groupe. Ce formant du chanteur est notamment dû à un abaissement du larynx et à un pharynx plus large. Il semble être responsable du timbre, du mordant dans la voix d'un chanteur ou d'un orateur formé.

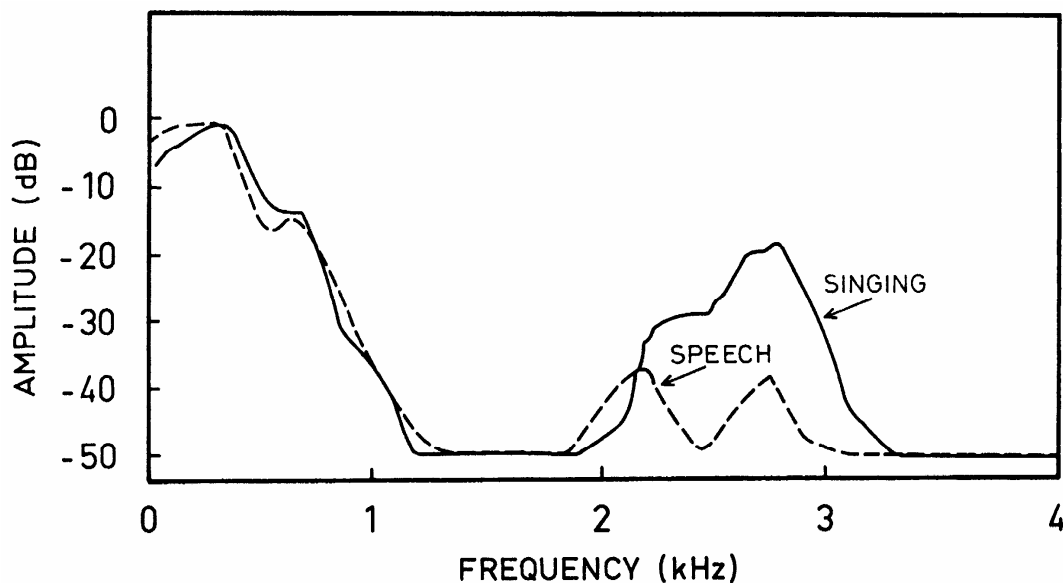


Figure 11 : Formants du locuteur et du chanteur décrit par Sundberg.

La fréquence de vibration correspond étroitement à notre perception de la hauteur (paramètre physique de la vibration) et de l'intonation (attribution perceptuelle du son). Cependant, la relation n'est pas linéaire. Le système auditif est beaucoup plus sensible à certains changements de fréquence qu'à d'autres. L'auditeur moyen est plus sensible aux changements de fréquence dans les zones basses. Par exemple, en augmentant le fondamental de 100 à 200 HZ, l'auditeur perçoit un

changement beaucoup plus important qu'une augmentation de 100 Hz dans la zone haute soit 3000 à 3100 Hz (Baken (1987)^[14]).

La capacité d'une voix entraînée d'être perçue au-dessus d'une chorale ou d'un orchestre dépend principalement de la présence de ce formant.

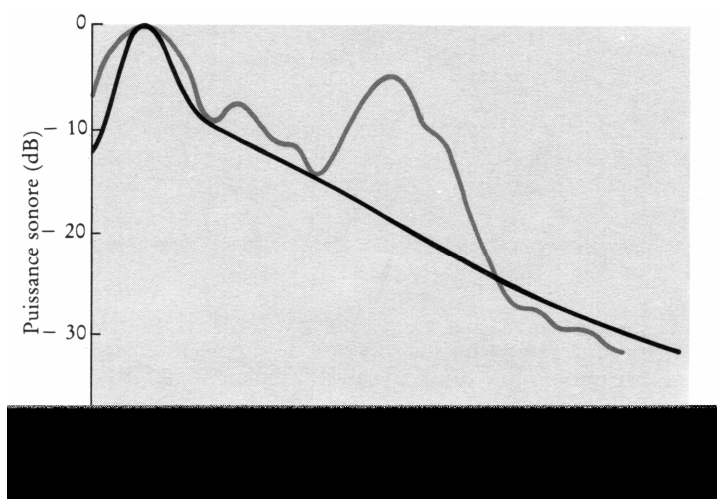


Figure 12 : Commentaires ci-dessous.

La figure 12 montre la distribution en fréquences de la puissance pour un orchestre et pour un ténor (Jussi Björling). Le niveau d'intensité du son orchestral baisse rapidement pour des fréquences supérieures à 500 Hz. La voix d'un chanteur présente un premier pic à environ 500 Hz, puis un autre à environ 2500 Hz ce qui lui donne un avantage sur l'orchestre de +/-13 dB (une intensité 20 fois supérieure). Pour cette valeur de la fréquence, le chanteur peut égaler l'orchestre, même s'il ne produit qu'un vingtième de sa puissance sonore. Ce formant contribue largement à différencier la couleur du timbre des différents registres vocaux. Chez les basses, il est observé à 2300 Hz, chez les barytons à 2600 Hz, chez les ténors à 2800 Hz, chez les mezzo-soprani à 2.900 Hz et chez les soprani à 3200Hz, mais peu observé chez les soprani aigus. Sundberg (1987)^[306] nuance cette observation. Il est difficile de déterminer avec précision les fréquences des formants lorsque la fréquence émise est élevée. Il stipule que si le formant est très élevé, il n'apparaîtra pas nécessairement, surtout si les filtres utilisés sont des passe-bandes étroits. Les harmoniques des sons vocaux sont espacés de manière équidistante sur l'axe des fréquences. Dans une voix

grave, les harmoniques sont nombreux et les formants peuvent dès lors être facilement identifiés. Sundberg^[in 269] souligne que cette configuration pourrait expliquer que les voix masculines ont été plus souvent analysées que les féminines.

Les fréquences des formants caractérisent les voyelles indépendamment de la hauteur. Ce n'est pas tout à fait vrai pour la voix d'un chanteur comme l'a montré Sundberg. D'une part, le chanteur s'efforce de produire un formant de chant élevé qui intensifie sa voix pour des valeurs de fréquence auxquelles les sons habituels d'un orchestre ont peu de puissance ; d'autre part, les soprani modifient leur formant lorsqu'elles chantent des notes très aiguës. Ces deux phénomènes intensifient le son et modifient sa qualité. La synthèse par ordinateur d'une voix chantée montre que si ces modifications n'ont pas lieu, la voix féminine ressemble à celle d'un enfant (Pierce (1984)^[233]). Les soprani abaissent la mandibule graduellement selon l'augmentation de la fréquence. La cavité buccale est agrandie. Comme, nous en avons déjà décrit les effets au paragraphe 2.2.2.1., l'intensité du 1^{er} formant est augmentée. Elles contractent également les commissures des lèvres. De cette façon, le tractus vocal est raccourci, et la fréquence de tous les formants diminue. Il peut également être raccourci par la position haute qu'il occupe dans la trachée. Certaines professionnelles emploient cette technique pour augmenter l'amplitude de leur 1^{er} formant. La forme de la langue change également avec la hauteur tonale, et ce principalement aux environs de 960 Hz.

Les études réalisées par Sundberg montrent que le principe du "tuning formant"⁹ s'appliquent au soprano et au ténor. Le bénéfice de cette technique (abaissement de la mâchoire, espace pharyngé élargi par le voile du palais) permet un gain sonore non négligeable puisqu'il peut atteindre 30 dB dans certains cas.

⁹ Le "**tuning formant**" consiste pour le chanteur à modifier les formants fréquentiels du tractus vocal afin d'augmenter certaines harmoniques de la source vocale. Ce principe existe probablement pour les autres tessitures excepté les basses.

1.2.2.5. L'Application à la voix chantée

L'augmentation de la pression du flux d'air contribue à augmenter le fondamental et l'intensité. L'explication de cet effet n'a pas été donnée de façon précise. Plusieurs expériences (in Miller(1986)^[200]) ont montré que le fondamental changeait de plus ou moins 4 à 5 Hz par cmH₂O pour une conversation ordinaire, ce qui n'est pas un changement de fréquence significatif. Par contre, pour un chanteur, cette modification est importante, mais grâce aux différents contrôles qu'il exerce sur les groupes musculaires, il parvient à en pondérer finement la dynamique.

La notion de registre est très controversée. Baken^[14] soulignait en 1987 qu'un consensus était loin d'être obtenu, même sur un point aussi fondamental que sa définition. Est-ce un attribut acoustique de la voix, un phénomène de perception, le produit d'un fonctionnement laryngé, une caractéristique du tractus vocal ? Hollien en 1974^[132] a tenté de mettre de l'ordre dans la confusion existante en définissant le registre comme le reflet d'un mode spécifique d'action laryngée, plutôt que l'effet d'une résonance supra-glottique. Chaque registre comprend une série ou une étendue fréquentielle qui peut être produite avec une qualité vocale plus ou moins identique. Baken impose trois conditions pour qu'il y ait registre laryngé :

Un registre laryngé devrait refléter un mode spécifique et distinct de l'action laryngée. Le registre modal est accepté implicitement comme la norme. Les registres les plus hauts ("loft") et les plus graves ("pulse") diffèrent du modal dans la forme et la tension des cordes vocales. Celles-ci sont massives en coupe frontale et l'étirement est réduit pour le registre pulsé, qui est atteint si toutes les tensions sont libérées, le fondamental est bas. Au contraire dans le registre loft, les cordes vocales sont extrêmement étirées et fines, le fondamental est élevé.

Miller (1986)^[200] estime, quant à lui, que les termes de registres modaux, loft et pulse "*ignorent les différences subtiles pouvant exister dans les écoles traditionnelles de chant en matière de définition nuancée de timbre.*" Les registres en chant sont prévisibles pour les différentes catégories vocales à quelques variations intra-individuelles près. Ces catégories sont déterminées en fonction des notes pivots marquant les limites du registres.

Pour les hommes, il rapporte quatre catégories de ténors (leggero, lirico, spinto, drammatico), deux catégories de barytons (lirico et drammatico) et deux catégories de basses (cantante et profondo). Le terme de falsetto dans le langage international du chant, précise-t-il, est galvaudé. Il définit l'imitation de la voix féminine telle que le réalise le chanteur masculin. En stroboscopie, on observe un manque continu d'accolement cordal. Le falsetto est confondu avec la voix de tête. *"En pédagogie, il ne s'agit pas de mélanger le falsetto à la voix de tête légitime de la voix d'homme, mais de développer la perception d'aisance associée au falsetto immédiatement avant le plein rapprochement des cordes tel qu'il se produit lors d'une émission vocale légitime"*.

Pour les femmes, il distingue trois catégories de soprani (coloratura, dramatico, leggero), la catégorie mezzo-soprano et la catégorie contralto. Le registre de "flageolet" rappelle le timbre de petites clochettes aiguës. Il est produit par une tension longitudinale des ligaments vocaux, une limitation de la masse vibrante des cordes et une forte pression sous-glottique. Un registre laryngé comprend une étendue fréquentielle.

L'étendue fréquentielle d'un registre laryngé donné recouvre l'étendue des autres registres et entraîne des chevauchements.

Dès lors, trois registres sont identifiés :

- **Le registre modal** désigne l'étendue des fréquences utilisées le plus habituellement en voix parlée (de 75 à 450 HZ pour les hommes ; de 130 à 520 Hz pour les femmes) et en voix chantée. L'appellation trouve son origine dans la terminologie statistique. Il signifiait "plus de valeur commune". Ce registre inclut des registres musicaux : poitrine, mixte et tête.
- **Le registre pulsé** (pulse) est produit dans l'étendue fréquentielle la plus basse (25 à 80 HZ pour les femmes et 20 à 45 HZ pour les hommes) L'émission laryngée est perçue comme un battement naturel. On parlera également de "vocal fry" ou de "glottal fry". Son synonyme musical est le "strobass". [Figure 13 A et B]
- **Le registre le plus haut** (loft) est produit dans l'étendue fréquentielle la plus haute (275 à 620 HZ pour les hommes et 490 à 1130 HZ pour les femmes). Le synonyme musical est "Falsetto". [Figure 13 A et B]

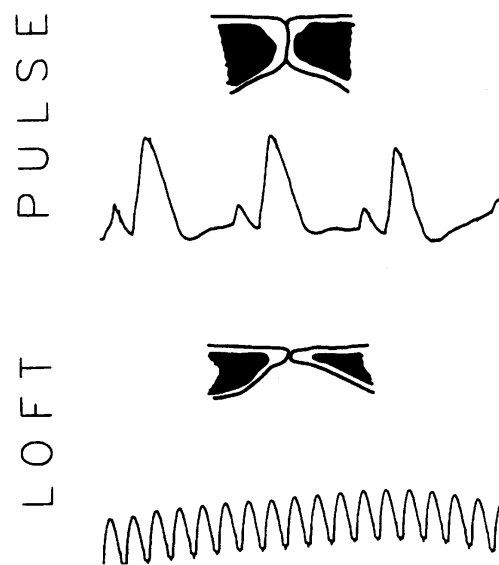


Figure 13 A : Cordes vocales en coupe frontale et ondes glottiques illustrant les registres "pulse" et "loft".

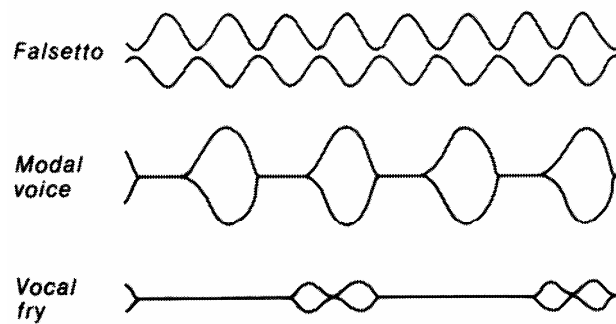


Figure 13 B : Modèle vibratoire normal des cordes vocales.

1.2.2.6. La vibration glottique

1.2.2.6.1. *L'irrégularité vibratoire*

"L'émission vocale est un phénomène quasi périodique. Quasi périodique car les bruits internes du corps, le mouvement irrégulier des impulsions électriques, des fluides et des cellules de l'organisme se traduisent par de petites irrégularités en fréquence, en amplitude et dans la forme de l'onde" (Giovanni (1998)^[98]).

Les pathologies cordales telles que nodule, polype, kyste, sulcus sont à l'origine d'irrégularités de la vibration cordale. Le mouvement périodique est altéré par cette perturbation. Les différences de vibration entre les deux cordes vocales s'accroissent nettement. Le flux aérien à travers la glotte n'est pas parfaitement régulier, il subit des mouvements aériens qui peuvent être en partie dus au hasard (on peut parler de "turbulence"). Le bouleversement ondulatoire se caractérise donc par de la turbulence et de l'irrégularité vibratoire.

Dans son workshop en 1995, Titze ^[329] clarifie certaines définitions concernant la périodicité du signal. Les termes de perturbation, fluctuation, variabilité, jitter, shimmer, tremor y sont expliqués. Nous présentons ici ces définitions.

Une **perturbation** est considérée comme étant un trouble mineur ou un changement temporaire d'un comportement attendu. Certaines petites irrégularités sont pour la plupart négligées.

Une **fluctuation** suggère une déviation plus sévère du modèle. Elle reflète une instabilité inhérente dans le système. Un système perturbé retourne d'habitude à la normale, un système fluctuant est incontrôlable et ne retrouve pas un état stable.

Un **tremor** ou un vibrato peut être décrit comme une fluctuation de l'amplitude et de la fréquence du fondamental. Il est son propre modèle plutôt qu'une petite déviation d'un modèle. Le tremor est une fluctuation à basse fréquence et basse amplitude. Son origine est d'habitude neurologique. Les tremors physiologiques dans le corps ont une cadence de 0 à 15 Hz. Si le vibrato peut être considéré comme un tremor

physiologique stabilisé dans les muscles laryngés, l'origine du vibrato vocal n'est pas complètement comprise.

Une **variabilité** est l'habileté de quelqu'un ou quelque chose à varier par accident ou par volonté.

Le **jitter** réfère à une perturbation à court terme (cycle à cycle) de la fréquence fondamentale de la voix. On doit cette appellation notamment à Lieberman et al (1982)^[182] qui a montré les courbes de paroles et a vu qu'aucune des deux périodes ne sont exactement identiques. La fréquence fondamentale paraît comme "sautillante" (jittery), d'où le terme de "jitter".

Le **shimmer** a été créé comme le pendant du jitter pour l'amplitude, soit la perturbation à court terme (cycle à cycle) de l'amplitude (Wendhal (1966)^[in 329]).

Déterminer la définition mathématique du shimmer et du jitter est une difficulté.

Titze^[329] rappelle les interrogations des mathématiciens : que signifie une perturbation à court terme?. Que peut-on adopter comme mesure de variabilité?

Pour la description des voix pathologiques, les termes de raucité, de souffle d'enrouement sont universellement utilisés.

La **raucité** (roughness) réfère à une qualité vocale irrégulière, cahoteuse. Elle résulte de l'irrégularité dans l'énergie contenue dans une bande critique du système auditif (Tehardt (1974)^[in 329]).

Le **souffle** (breath) est la qualité vocale qui contient du souffle expiratoire durant la phonation. Le signal dû à la turbulence glottique de l'air comporte une composante acoustique significative.

L'**enrouement** (hoarseness) désigne la combinaison de raucité et de souffle.

Ces termes, précise Titze, qui n'ont pas de définitions mathématiques peuvent pour la plupart être évalués sur un plan psychophysique.

1.2.2.6.2. Les mesures de perturbation : le Jitter, le shimmer, le rapport signal/bruit.

Une mesure de perturbation est une valeur efficace de la perturbation globale du contour du cycle. Les ondes glottiques sont imparfaites non seulement à cause des

turbulences du flux mais aussi à cause des perturbations de fréquence et d'amplitude. (Sataloff (1997)^[269]). Cette perturbation de cycle à cycle est appelée dans la littérature, jitter pour la fréquence et shimmer pour l'amplitude. Les différences moyennes sont très petites, de l'ordre de 40 micro secondes pour la période et de 0.4 dB pour l'amplitude.

Un des problèmes majeurs de la physique acoustique concerne la validité de ces mesures de perturbation. Cette difficulté augmente avec le degré d'apériodicité de l'échantillon vocal. La fiabilité de la mesure de perturbation est inversement corrélée avec le degré de sévérité de la dysphonie (Bielamowisc et al (1996)^[21], Titze et al (1993)^[326]).

L'erreur standard (produit de la statistique) est la valeur absolue (carré) de la différence entre le résultat d'une mesure et le résultat qui devrait être obtenu s'il n'y avait aucune irrégularité (soit la moyenne). L'erreur standard, qui prend le nom de coefficient, correspond à la variance dont la racine carrée est l'écart-type.

Le **jitter** est une mesure instantanée qui est un indice d'apériodicité à court terme du signal. Il peut être exprimé en valeur absolue (milliseconde ou hertz) ou en valeur relative. Il est alors le pourcentage moyen de variation de la fréquence d'un cycle à l'autre (Giovanni (1995)^[93]). Cependant, exprimé en valeur absolue, il est influencé par la fréquence du fondamental moyen, les mesures relatives sont dès lors plus couramment employées. Cette mesure de variabilité en fréquence ne prend pas en considération les changements volontaires de la fréquence fondamentale. Si le système phonatoire est dans un état de stabilité mécanique parfait et idéal, il n'y aura strictement aucune différence dans les périodes du fondamental lorsque le locuteur changera volontairement d'intonation.

La relation jitter-fréquence n'est pas linéaire. Orlikoff et al (1991)^[223] a rapporté des valeurs de jitter significativement différentes entre hommes et femmes, ce que confirme Giovanni^[95] par une étude sur 66 sujets normaux ($p=0.044$).

Tableau 13 : Valeurs de jitter rapportées par Giovanni et al (1996)^[95].

Sujets	Tâches	Mesures	Valeurs H	Valeurs F	P
33H/33F	[a]	Fo	126.6 (18.6)	223.6 (33.2)	P=0.000
		Jitter	0.7 (0.3)	0.5 (0.1)	P=0.044

Légende : H : hommes, F : femmes.

Le **shimmer** est une mesure instantanée, indice d'apériodicité à court terme du signal. Il peut être exprimé en valeur absolue (décibel) ou en valeur relative. Il est alors le pourcentage moyen de la variation d'amplitude d'un cycle à l'autre.

Les mesures du jitter et du shimmer sont employées comme index de l'habileté phonatoire et sont sensibles aux inférences liées à l'âge (Linville ^[in 326], Wilcox et al (1980)^[345]), à la croissance et au développement chez l'enfant (Glaze et al (1988)^[99]), au dysfonctionnement neurologique (Ramig et al (1983)^[243]) et à la chirurgie (Hall, Newman) in Ferrand et al (1996)^[80].

1.2.2.6.3. La validité des mesures

- En fonction des appareils.

Les comparaisons des valeurs de jitter d'un laboratoire à un autre sont difficiles et ce, malgré la connaissance des méthodes d'extraction de la fréquence fondamentale et des algorithmes de calcul. Titze et al (1993)^[326] rapportent les valeurs suivantes :

Tableau 14 : Valeurs de jitter en fonction des méthodes de calcul d'extraction de la fréquence fondamentale ^[326].

Méthodes	Valeurs homme (%)	Valeurs femmes (%)
Wave- Form matching ¹⁰	0.07- 0.7	0.08- 0.8
Peak Picking ¹¹	0.13- 1.8	0.23- 1.19
Zéro- crossing ¹²	0.12- 0.78	0.13- 0.67

Karnell et al (1991)^[156] ont démontré qu'il pouvait exister d'importantes variations en fonction de l'appareillage utilisé. Ils rapportent des valeurs de jitter

¹⁰ **Le Wave-Form matching** : méthode d'extraction de la fréquence fondamentale qui consiste à comparer le signal par rapport à lui-même en déplaçant la fenêtre d'analyse (calcul d'auto-corrélation).

¹¹ **Le Peak Picking** : méthode manuelle d'extraction de la fréquence fondamentale qui consiste à déterminer les périodes du signal filtré.

¹² **Le Zéro-crossing** : méthode automatique d'extraction de la fréquence fondamentale qui consiste dans un premier temps à filtrer le signal pour éliminer les harmoniques et dans un second temps à décompter le nombre de passage par 0 du signal et ensuite à le diviser par deux.

significativement différentes entre le visipitch 6097/IBM PC et deux autres systèmes mini-computers. Il attribue cette diversité au filtre utilisé. Les valeurs obtenues sur les deux autres systèmes montrent par contre une corrélation de 0.79 à 0.89 avec une différence moyenne inférieure à 4%. Perry et al (1996)^[230] ne se satisfont pas de l'hypothèse du choix des filtres pour justifier la différence obtenue. Ils comparent à leur tour les valeurs de jitter obtenues par méthode manuelle en employant la formule de Koike et le Visi-Pitch. Les résultats indiquent que les différences de valeurs de jitter obtenues par les deux systèmes, sont petites surtout lorsque l'utilisateur emploie les filtres conseillés par le fabricant. Perry cherche à expliquer la divergence de résultats et d'hypothèse entre Karnell et al et eux. Karnell et al ont capturé les informations sur enregistreur analogique et les ont envoyés dans 3 laboratoires différents. Deux de ces laboratoires ont diffusé le signal avec un enregistreur différent de l'original. Perry et al soulignent que cela peut produire des erreurs instrumentales qui influeraient sur la précision de l'estimation du jitter. Ce fait a été démontré par Doherty et al en 1988^[67]. Il ajoute que les changements de vitesse de déroulement de la bande magnétique, la tête de l'enregistrement et d'éventuels problèmes de contact sont autant de facteurs qui influent sur la mesure. Bough et al (1996)^[29] comparent le laryngographe de Fourcin et le Visi-Pitch. Les valeurs de fréquence fondamentale sont fiables dans les multiples échantillons effectués sur plusieurs jours ; par contre, les valeurs de jitter et shimmer sont hautement variables (pour les femmes 0.25 versus 0.70 ; pour les hommes 0.55 versus 0.60).

Tableau 15 : Valeurs de jitter selon différents auteurs.

Auteurs	Années	Jitter (%)	Valeurs H	Valeurs F
Holien et al ^[131]	1973	0.3- 1.3	Sexe non précisé.	
Horii ^[138]	1979	0.7		
Murphy et al ^[211]	1980	0.8- 1.5		
Horii ^[140]	1982	0.66		
Sorensen et al ^[291]	1983	0.71		
Baken ^[15]	1987	0.37		
Ramiget al ^[243]	1983	0.4- 0.9		
Klingholz et al ^[166]	1985	0.3- 1.6		
Jafari et al ^[153]	1993		0.20- 0.22	0.24- 0.34
Ferrand et al ^[80]	1995			0.682
Bough et al ^[29]	1996		0.25	0.55

A la vue des divergences observées, Giovanni (1998)^[98] propose certaines conditions dans l'emploi clinique du jitter. En accord avec Schoentgen (1989)^[277], il déclare que les mesures devraient être effectuées sur la partie stable d'une voyelle tenue. Et en accord avec Karnell et al (1991)^[156], que la longueur de la fenêtre d'analyse ainsi que sa position devraient être plus largement considérées. Concernant les méthodes d'extraction, la mesure du jitter est directement dépendante du système de détection de la période (Titze (1995)^[329], Deem et al (1989)^[65]). Si la période est difficile à détecter, la mesure du jitter n'en sera que moins fiable. Concernant les comparaisons, les résultats sont difficilement reproductibles d'un laboratoire à l'autre (Karnell et al (1995)^[157], Bough et al (1996)^[29])

La perturbation de l'amplitude est généralement mesurée par le shimmer. A l'instar de la perturbation fréquentielle, les calculs de shimmer servent à quantifier l'instabilité à courte de durée du signal vocal. Cette mesure a été créée pour être le pendant du jitter dans le domaine de l'intensité (Giovanni (1998)^[98]).

Titze et al (1987)^[320] comparent les mesures spécifiques de jitter % et de shimmer % calculées immédiatement par un logiciel d'analyse et ces mêmes mesures via un enregistreur. Les valeurs de jitter restent plus ou moins stables (0.14% versus 0.19%). Ce n'est pas le cas pour les mesures de shimmer (0.51% versus 0.98%). Doherty et al (1988)^[67] analysent cette variabilité au départ de 4 types d'enregistreurs différents. Ils confirment un haut degré de variabilité pour le shimmer, dont la mesure est particulièrement sensible aux différents calibrages des magnétophones, enregistreurs numériques et vidéo scopes. Gelfer et al (1995)^[88], aussi, rapportent des variations significatives pour le shimmer (1.29% versus 1.72%).

- En fonction de la pathologie.

Les méthodes temporelles d'extraction de la fréquence fondamentale telles que le "peak picking", le "wave-form matching" ou encore le "zero-crossing", présentent un certain nombre de difficultés dues à l'apériodicité du signal. La détection de la période est d'autant plus difficile que le signal est instable. La vibration des cordes peut être simplement périodique, a-périodique, quasi périodique et aléatoire. Dans le premier cas, une période correspond à un cycle ; dans le second, les cycles ne sont pas de

même longueur mais des séquences identiques peuvent se répéter, on parlera de quasi-périodicité., Dans le troisième cas, chacune des cordes vibre à une fréquence différente. On parle de diplophonie. Dans le quatrième cas, la vibration peut être aléatoire, les cycles ne se reproduisant jamais.

1.2.2.6.4. Les corrélations des mesures de perturbation avec la perception

Tableau 16 : Corrélations entre paramètres acoustiques et paramètres perceptuelles selon différents auteurs.

Auteurs	Années	Paramètres Acoustiques	Qualités perceptuelles	Corrélations
Dealet al ^[50]	1978	Jitter	Raucité	0.03-0.69
Kojima et al ^[167]	1980	HNR	Hoarse	-0.86
Yumoto et al ^[359]	1984	HNR	Hoarseness	-0.80
		Jitter	Hoarseness	<u>0.71</u>
Wolfe et al ^[350]	1987	Jitter	Raucité	<u>0.51</u>
Eskenazi et al ^[77]	1990	HNR	Raucité	>0.80
		Jitter	Souffle	<u>>0.80</u>
Dejonckere et al ^[61]	1993	Jitter	Raucité	<u>0.68</u>
		Shimmer	Grade	0.73
		HNR	Grade	0.64
		Shimmer	Souffle	0.63
de Krom ^[55]	1995	Mesures de perturbation fréquentielle	Raucité	<u>0.66 – 0.76</u>
			Enrouement	<u>0.58 – 0.68</u>
Wolfe et al ^[352]	1995	Shimmer	Gravité de la dysphonie	0.54
		HNR	Gravité de la dysphonie	-0.32
Martin et al ^[189]	1995	Jitter	Raucité	<u>-0.85</u>
		Shimmer	Raucité	0.66
		HNR	Raucité	0.50
Wolfe et al ^[353]	1997	Jitter	Voix soufflée	<u>0.63</u>
			Voix Serrée	<u>0.58</u>
		Shimmer	Voix Soufflée	0.52
			Voix enrouée	0.72
			Voix serrée	0.58
		HNR	Voix soufflée	-0.51
			Voix enrouée	-0.76
		Jitter SD	Voix soufflée	0.59
		Shimmer SD	Voix soufflée	0.55
Wolfe et al ^[354]	1997	Jitter	Voix soufflée	<u>0.61</u>
			Voix enrouée	0.57
			Voix rauque	0.57
		Shimmer	Voix soufflée	0.67
			Voix enrouée	0.59
			Voix rauque	0.46
		HNR	Voix soufflée	0.68
			Voix enrouée	0.63
			Voix rauque	0.54
Millet et al ^[201]	1998	Jitter%	Raucité	0.37
		Shimmer %	Souffle	0.49

Légende : SD : déviation standard ; HNR : "harmonics/noise ratio" ou rapport harmoniques/bruit.

Certaines études ont montré que les mesures acoustiques telles le jitter, le shimmer ou encore le rapport signal/bruit sont associées avec des qualités perceptuelles telles que le souffle, la raucité ou l'enrouement. Cependant, ces corrélations varient largement d'un auteur à l'autre.

Concernant le **jitter**, déjà en 1978, Deal et al ^[50] montraient que la corrélation raucité/ jitter chez 20 sujets dysphoniques émettant 5 voyelles différentes pouvait varier de 0.03 à 0.69. Yumoto et al ^[359] rapportent une corrélation de 0.71 entre l'enrouement et le jitter pour 87 sujets dont le degré d'enrouement varie de normal à sévère. Wolfe et al ^[350] trouvaient une corrélation modérée de 0.51 entre la raucité et le jitter pour les voyelles [a], [i], produites par 51 sujets dysphoniques suite à une large variété de problèmes vocaux. En 1997, Wolfe et al ^[353] montrent des corrélations entre le jitter et la voix soufflée de 0.63, ainsi qu'entre le jitter et la voix serrée de 0.58. Ils affinent ces corrélations et rapportent dans une autre étude une corrélation jitter/souffle de 0.61 ; jitter/enrouement de 0.57 et jitter/raucité de 0.57. de Krom (1995)^[55], quant à lui, rapporte des corrélations entre les mesures de perturbations fréquentielles et la raucité variant de 0.66 à 0.76, ainsi que des corrélations entre ces mêmes mesures et l'enrouement de 0.58 à 0.68. Eskenazi et al (1990)^[77], quant à eux, trouvent que le jitter% est un des meilleurs prédicteurs de sévérité de la raucité et du souffle avec des corrélations supérieures à 0.80. Martin et al ^[189] rapportent une corrélation jitter/raucité de 0.50, Dejonckere et al ^[61] de 0.68. Millet et al (1998)^[201] rapportent des valeurs de corrélation sur une population de 65 sujets ayant une voix pathologique. Deux juges évaluent l'émission de la voyelle [a] à l'aide de l'échelle G.R.B.A.S.. La corrélation jitter%/souffle est de 0.37. Rabinov et al (1995)^[241] rapportent que les auditeurs sur-évaluent la gravité de la dysphonie tandis que les valeurs de jitter sont basses. Hillebrand stipule qu'au plus la fréquence fondamentale augmente, au plus le jitter diminue.

Concernant le **shimmer**, Dejonckere et al (1993)^[61] rapportent une corrélation shimmer/grade de 0.73, et shimmer/souffle de 0.63. Ce dernier résultat est corroboré par Wolfe et al (1997)^[353] qui rapportent des corrélations de 0.67 pour shimmer/souffle, 0.59 pour shimmer/enrouement et 0.46 pour shimmer/raucité. Dans une autre de ces études, les résultats s'inversent. Les corrélations sont meilleures pour shimmer/voix enrouées (0.72) tandis que shimmer/voix soufflées montre une

corrélation de 0.52. Martin et al (1995)^[189] trouvent, quant à lui, une corrélation plus élevée pour shimmer/raucité : 0.66. Millet et al (1998)^[201] rapportent une corrélation Shimmer %/raucité de 0.49.

Concernant le rapport signal/ bruit, Isshiki (in Cornut (1991) ^[46]) a montré que le degré de raucité dépendait surtout du rapport entre la composante de bruit et la composante d'harmonique du spectre vocal. de Krom (1994)^[54] partage cet avis. Il dit que le HNR semble être un des paramètres qui convient pour établir un lien entre les aspects physiologiques de la production vocale et la perception de la voix. Pour lui, le degré de bruit spectral est lié à la qualité de l'émission vocale. Kojima et al (1980)^[167], Yumoto et al (1984)^[359], Martin et al (1995)^[189], Eskenazy et al (1990)^[77] et de Krom (1995)^[61] trouvent tous qu'effectivement HNR est un prédicteur significatif des échantillons vocaux évalués perceptuellement comme rauques. Ils trouvent des corrélations de 0.80 et plus. Dejonckere et al (1993)^[61] trouvent une corrélation HNR/ Grade (échelle G.R.B.A.S.) de 0.64. Wolfe et al, quant à eux, rapportent de meilleures corrélations pour HNR/ souffle (0.68). Les corrélations HNR/enrouement et HNR/ raucité sont plus basses, respectivement 0.63 et 0.54. Dans une autre étude, les auteurs rapportent des corrélations HNR/voix soufflée de -0.51, et HNR/voix enrouée de -0.76. La corrélation qu'ils observent entre HNR et la gravité de la dysphonie est faible (-0.32).

Au vu de ces résultats, il semble difficile de dégager des corrélations spécifiques entre les paramètres objectifs et les qualités perceptuelles. Le rapport signal/ bruit semble être néanmoins un prédicteur de la sévérité de la dysphonie, les résultats des corrélations étant généralement plus élevés que pour le jitter ou le shimmer. Le jitter semble autant lié au souffle qu'à la raucité avec des corrélations variant de 0.55 à 0.80 pour le souffle et 0.51 à 0.85 pour la raucité. Le shimmer présente les mêmes caractéristiques avec des corrélations variant de 0.49 à 0.63 pour le souffle et 0.46 à 0.66 pour la raucité.

Une des raisons de ces résultats divergents serait que la raucité, bien que perceptuellement distincte du souffle et de l'enrouement, est une qualité pathologique vocale qui varie considérablement d'une étude à l'autre. Pour expliquer ces divergences, Lending suppose que les facteurs physiologiques contribuant à la dysphonie sont nombreux et variés. Un autre possibilité serait que les mesures

acoustiques sous des appellations similaires diffèrent sur base du design ou du mode d'exécution. Une troisième possibilité serait que la prédiction acoustique de la sévérité de la dysphonie puisse découler d'une erreur de mesure. Plusieurs mesures automatisées sont sensibles aux erreurs dans la "démarcation" de la période de la hauteur tonale.

Il est probable donc que les différentes mesures acoustiques prédisent la sévérité de la dysphonie avec un succès variable dépendant du type de voix considéré.

Le calcul du jitter dépend du cycle périodique. Si le signal vocal est trop irrégulier pour générer un cycle périodique à l'intérieur d'une chaîne déterminée, le jitter ne peut être obtenu. Dans l'étude de Plant et al (1997)^[234] sur 16 sujets ayant une paralysie unilatérale, les auditeurs évaluent les voix sur un seul paramètre : la qualité globale de la voix en pré et en post opératoire. Les valeurs ne sont pas obtenues pour 8 échantillons. Pour les autres, la corrélation jitter/grade est très faible 0.07. Les troubles vocaux dus à une paralysie vocale unilatérale reste un défi pour les mesures acoustiques de la qualité vocale. La turbulence au niveau de la glotte produit des composantes non voisées accompagnées de larges variations de fréquence sur de petites périodes. Dans ce cas, les auteurs font remarquer que le calcul de la fréquence basée sur la technique du pic à pic est inefficace vu le bruit important présent dans le signal vocal. Ceci corrobore les dires de Rabinov et al (1995)^[241] : le jitter serait donc plus approprié pour l'évaluation de petits changements, pour les voix normales mais pas pour les troubles vocaux sévères. Les études rapportées ci-dessus mentionnent que les sujets observés souffrent de pathologies variées, voire très variées, entraînant un degré de dysphonie qui s'étend de léger à sévère ; ces considérations sur le jitter doivent entrer en ligne de compte et peuvent expliquer également la divergence des résultats obtenus.

chapitre 2

La vidéostroboscopie

2.1. LA VIDÉOSTROBOSCOPIE

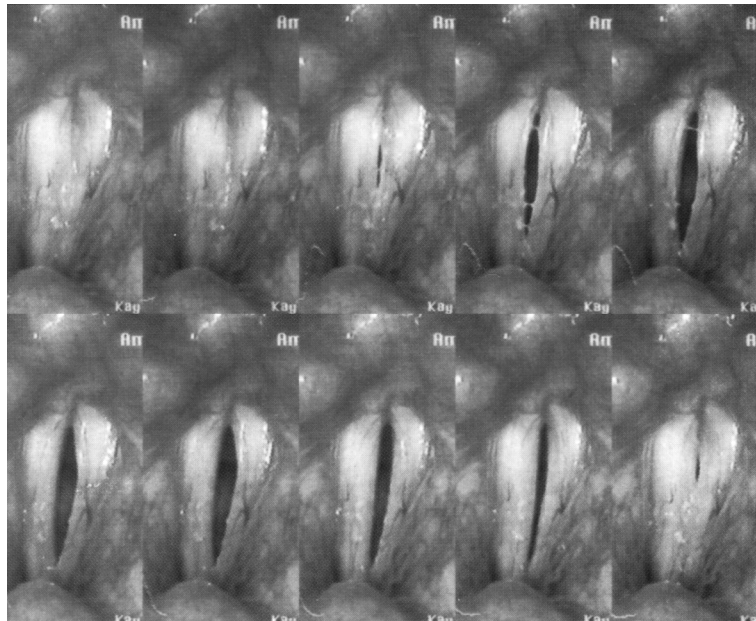


Figure 14 : Cycle vibratoire, images digitalisées sur écran.

2.1.1. Description de la technique.

Observer la vibration des cordes vocales en temps réel est impossible. Les cordes vocales vibrent à une vitesse de 100 cycles par seconde chez l'homme et 250 cycles chez la femme. Cette fréquence est trop rapide pour être observée à l'œil nu, mais la cinématographie à haute vitesse et la stroboscopie peuvent rendre compte de ces mouvements complexes et rapides.

La vidéostroboscopie, procédé fondé sur une illusion optique, est la méthode la plus fréquemment employée de nos jours pour l'examen du larynx et l'étude de la vibration des cordes vocales.

Le stroboscope, lui, est basé sur la LOI de TALBOT. Celle-ci développe le phénomène de "persistance de la vision" : l'image perçue par la rétine persiste 0.2 secondes. L'œil perçoit donc 5 images par seconde.

Du fait que l'œil ne peut percevoir que 5 images par seconde, les séquences d'images produites à un intervalle plus court que 0.2 secondes persistent sur la rétine et les images fusionnent pour produire une illusion optique de mouvement apparent. Il s'y ajoute un phénomène de correspondance. Ceci signifie que lorsque le système visuel perçoit une série rapide d'images fixes [figure 15], le cerveau est capable de combler les vides entre les images et d'imaginer qu'il voit un objet dans un mouvement continu pour autant que les images successives ne diffèrent que légèrement les unes des autres. Si les images diffèrent trop, elles apparaissent secouées comme à l'époque des premiers cinématographes [figure 16].

Le matériel employé pour la stroboscopie comporte un capteur de vibration, une source lumineuse, une unité de contrôle électronique et une pédale de commande au pied.

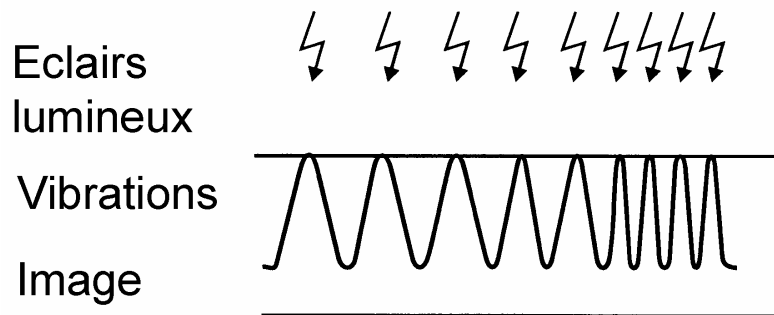


Figure 15 : Principe de la stroboscopie. L'image fixe.

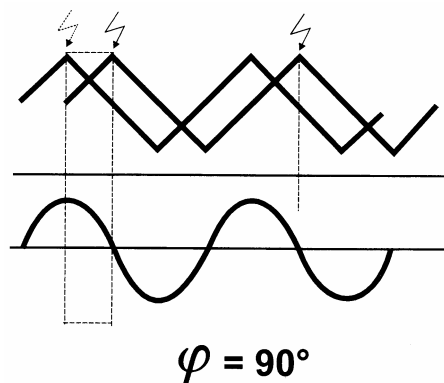


Figure 16 : Principe de la stroboscopie. Le déphasage.

Les stroboscopes électroniques apparaissent au début des années 1950. Les vibrations laryngées sont transmises au stroboscope par la membrane du capteur de vibration placé sur le cou du sujet. Les vibrations sont transformées en pulsations électroniques et le fondamental produit est transmis à une lampe xénon. Celle-ci produit des flashes de lumière blanche émis à la même fréquence que la vibration cordale. Le flash de fréquence peut être augmenté de 2 à 4 Hz pour donner l'impression d'un mouvement. Une pédale permet de passer d'un mode fixe à un mode ralenti [Figure 17].

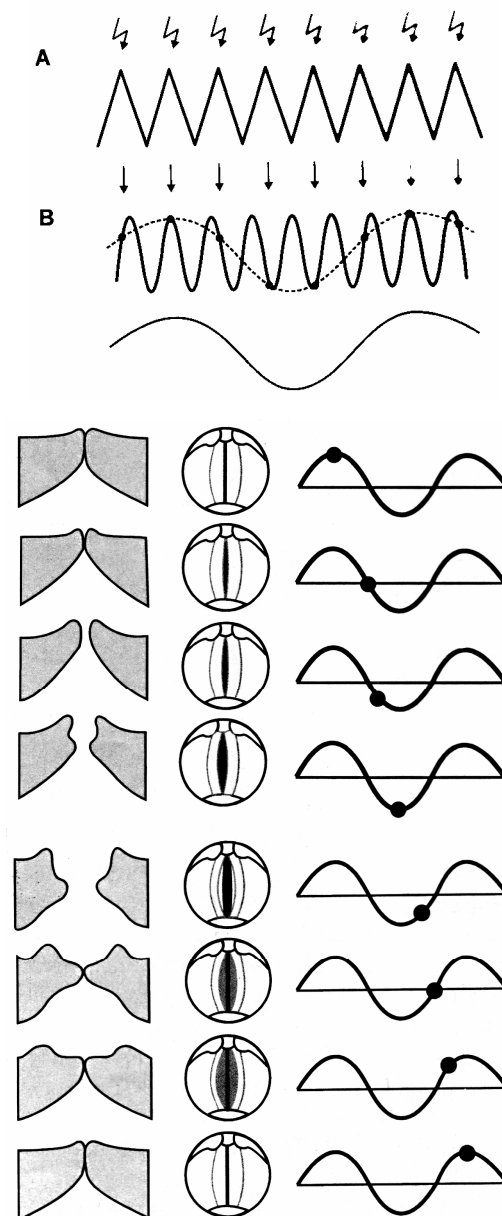


Figure 17 : Principe de la stroboscopie, l'impression de ralenti.

L'équipement est amélioré en fonction de l'usage clinique. L'optique laryngologique rigide de Barth (1976)^[17] connectée au stroboscope fait partie de la pratique quotidienne de l'O.R.L.. Yoshida (1979)^[358] utilise l'enregistrement des images laryngées grâce à une caméra connectée au système optique. Le faisceau lumineux peut être dirigé sur le larynx de plusieurs manières, soit par la laryngoscopie rigide, soit par la fibre optique, soit encore par un microscope d'opération. Le fibroscope et l'optique rigide sont fréquemment utilisés et de manière complémentaire. Ces deux instruments sont complémentaires. L'optique rigide permet une vue générale amplifiée du larynx. Les changements subtils produits par la muqueuse peuvent être observés. L'axe visuel et l'axe de lumière sont alignés. Cependant, l'inconvénient majeur de cette méthode est le réflexe vélopharyngé. Le sujet examiné ne peut produire que des voyelles. Les conditions d'émission ne sont pas optimales (position anormale de la tête, langue tirée). Si on emploie le fibroscope seul (naso-fibroscope), le sujet peut être examiné sans réflexe. L'émission phonatoire peut-être parlée ou chantée. Cependant, il ne permet qu'une observation plus globale et l'image est de qualité plus petite et plus pauvre. Les fibroscopes perdent entre 60 et 90% de lumière disponible. Une petite quantité de lumière est produite et, par conséquent, une source lumineuse plus puissante est requise. Une comparaison a été effectuée par Costamagna (1990)^[47]. A établi une comparaison entre l'optique rigide et le naso-fibroscope.

L'avantage du mode stationnaire, c'est que les images sont extrêmement claires et supérieures à celles obtenues en fixant l'image ou mouvements lents en progressant pas à pas.

L'examen de la vibration lente donne une image générale de la vibration en fonction de l'intensité (en moyenne, l'examen se fait à 60-70 dB) et du registre. Dans un premier temps, la phonation est étudiée dans le registre de parole ou un registre intermédiaire (en moyenne : 220 à 250 Hz pour les femmes et 110 à 150 Hz pour les hommes). La vibration laryngée est étudiée si possible durant l'émission des sons dans le registre de poitrine (basses fréquences) et le registre de tête (hautes fréquences). Pour stabiliser la fréquence fondamentale, il est nécessaire de prolonger l'examen d'au moins une seconde. L'étude des variations de la vibration comme une fonction de la hauteur tonale est importante chez les chanteurs et les professionnels de la voix. Durant les hautes fréquences, la mise sous tension des cordes vocales aidera à révéler les lésions intra et sous muqueuse telles que le kyste épidermoïde. La

vibration symétrique entre les 2 cordes vocales (symétrie d'amplitude et symétrie de phase) peut aussi être observée. La vibration est symétrique en amplitude quand le déplacement est identique pour les deux cordes vocales. La vibration est symétrique en phase quand les moments de fermeture et d'ouverture maximaux sont synchrones pour les deux cordes. La symétrie d'amplitude peut être affectée pendant que la symétrie de phase ne l'est pas et vice-versa. Il y a des différences dans les propriétés mécaniques des cordes vocales (élasticité, viscosité, tension, volume, masse, forme, position,...) lesquelles donnent une élévation du mouvement d'asymétrie. La périodicité ou la régularité peuvent être examinées. La régularité concerne les cycles successifs d'une même corde vocale. Cette caractéristique est mieux vue en mode stationnaire. En effet, si la vibration est réellement régulière, l'image du bord libre de la corde vocale apparaît statique dans le mode stationnaire. Si, au contraire, il y a plusieurs apériodicités ou irrégularités, alors l'O.R.L. observe un tremblement du bord libre. Pour obtenir une vibration périodique, les propriétés mécaniques des cordes vocales et la pression respiratoire sous-glottique doivent être en équilibre. Celui-ci peut être perturbé par les différences dans les propriétés mécaniques des cordes vocales soit l'incapacité à maintenir un tonus constant des muscles vocaux et à maintenir une pression sous-glottique constante.

La stroboscopie est indiquée pour vérifier la qualité de la fermeture glottique, les cordes vocales apparaissant floues en condition de lumière continue. La fermeture glottique est définie comme normale quand les cordes vocales sont en contact ou approximativement fermées. Cela correspond à 0° cyc le vibratoire.

La durée de la fermeture glottique augmente lorsque que la hauteur diminue et que l'intensité vocale augmente¹³.

La position de la fermeture de la glotte peut aussi être évaluée : elle est généralement médiane, mais peut également être para-médiane dans le cas, par exemple, d'une paralysie récurrente unilatérale avec une compensation controlatérale. La fermeture incomplète de la glotte peut résulter d'une adduction incomplète des cordes vocales, d'une irrégularité du bord de l'une ou des deux cordes vocales, de plusieurs obstacles entre les cordes vocales et de la rigidité du bord de l'une ou des deux cordes vocales. La stroboscopie s'emploie aussi pour calculer le temps de variation des phases du cycle vibratoire. Ainsi, peuvent être mesurés :

- le temps de fermeture glottique,
- le temps d'ouverture glottique,
- le quotient d'ouverture défini comme la vitesse du quotient est égale à la durée du temps d'ouverture sur le temps de fermeture.
- la vitesse du quotient ou l'index de vitesse correspondant au quotient de vitesse – 1 sur le quotient de vitesse +1.
- l'amplitude définie comme l'étendue du déplacement latéral de la corde vocale ; sa mesure incluant le tracé de l'écart latéral ou abduction des cordes vocales suivi du mouvement de rapprochement simultané des cordes vocales vers la ligne médiane ou adduction. Physiologiquement, ce mouvement correspond à une vibration musculo-ligamentaire basique.

L'amplitude peut dépendre des variations physiologiques : son accroissement est lié à l'augmentation de l'intensité vocale, sa diminution à l'augmentation de la fréquence fondamentale, une disparition de l'amplitude peut même être observée chez les falsettistes¹⁴.

L'ondulation de la muqueuse résulte de son déplacement en vague sur le muscle vocal sous-jacent. C'est en fait, un déplacement qui est additionné à une vibration musculaire basique. L'ondulation de la muqueuse se produit sous l'apparence d'une forme plane, et verticale (de bas en haut) avec un déphasage entre le plan supérieur et le plan inférieur du bord libre de la corde vocale.

L'examineur recourt généralement à un protocole de base. Une version simplifiée de celui d'Hirano et Bless (1993) ^[127] est employée dans la pratique quotidienne des O.R.L. du Belgium Study Group Voice Disorders (B.S.G.V.D.) [Annexe 2].

¹³ Bless et al (1987)^[24] ont proposé une classification des formes incomplètes de fermeture glottique : fuite longitudinale, fuite postérieure, fuite antérieure, glotte "hour-glass", glotte ovalaire et glotte irrégulière.

¹⁴ Les falsettistes sont également appelés Hautes-contres. Cette voix masculines se situe dans le registre aigu du ténor.

2.1.2. Les limites de la vidéostroboscopie.

Il s'agit de toujours garder à l'esprit que le cycle vibratoire observé est une composition d'images à partir de plusieurs cycles, et que celles-ci sont des illusions d'optique. Si le phénomène est irrégulier et peu fréquent, il peut échapper à l'examen. Seules des images en lumière continue permettent d'observer la totalité du phénomène.

En réponse à cette difficulté est née la cinématographie à haute vitesse. Elle peut capturer jusqu'à 10.000 images par seconde. Cette technique reste cependant l'apanage de quelques laboratoires. Son équipement est lourd, et des difficultés subsistent à l'enregistrement.

Hirose (1998)^[128] a développé un système à haute vitesse, informatisé. Les images obtenues restent pauvres en qualité et en détail. Les avancées technologiques pallieront ces difficultés.

2.1.3. L'examen du sujet.

Sulter et al (1996)^[302] proposent de déterminer l'influence des facteurs de sexe, d'entraînement vocal, d'intensité sonore, de hauteur tonale, et d'âge sur le fonctionnement vocal de 214 sujets appariés en fonction du sexe et de l'état d'entraînement vocal. Trois juges évaluent le fonctionnement vocal sur base de protocoles vocaux standardisés comprenant **l'aspect laryngé** (le rapport larynx/pharynx, la forme de l'épiglotte, l'asymétrie aryénoïdienne, les ajustements compensatoires, l'épaisseur, la largeur, la longueur, et l'élasticité des cordes vocales) et **le fonctionnement glottique** (l'amplitude de la course, la durée, le pourcentage, et le type de fermeture glottique, les différences de phases, et la localisation de la fuite glottique). Les sujets prononcent un [i]. Cette voyelle permet une meilleure visibilité du larynx grâce à une position antérieure maximale de la partie supérieure de l'épiglotte. L'accord inter-juges est élevé (α : 0.65), excepté pour les échelles concernant l'épaisseur (α : 0.45), l'élasticité des cordes vocales (α : 0.28), et l'évaluation du rapport larynx/pharynx (α : 0.43). De légères différences sont observées entre les sujets non-entraînés et entraînés. Ces derniers ont plus souvent une fermeture glottique complète et montrent une différence de phase entre les cordes vocales de

type latéral. Des différences dans l'apparence laryngée sont établies entre les sexes masculin et féminin. Les caractéristiques laryngées et glottiques féminines ne sont pas comparées simplement aux caractéristiques masculines transposées d'une octave. Les cordes vocales masculines sont plus épaisses, et plus tendues. Avec l'âge, les hommes montrent une augmentation de la complexité de fermeture. Les sujets plus âgés dont la structure des cordes vocales est atrophiée seraient obligés de les accoler plus intensément. Chez les femmes, l'âge entraîne une augmentation de la durée de fermeture. Il y aurait moins de pression sous-glottique. Une différence de phase horizontale est observée plus fréquemment chez les femmes tandis qu'une différence de phase verticale l'est plus souvent chez les hommes. Les hommes montrent un temps de fermeture plus court, ainsi qu'un degré d'accolement plus important.

Les différences dans l'apparence des cordes vocales et leur mode de vibration sont significatives en fonction de la hauteur tonale, de l'intensité et de l'âge.

Elias et al (1997)^[74] soulignent également l'importance qu'il y a à établir un protocole de base strobovidéolaryngoscopique dans l'évaluation des performances des professionnels de la voix. Les auteurs y contribuent en réalisant une revue sur 65 professionnels du chant en bonne santé. Ils répertorient 58% d'anormalités classées en six entités cliniques, les irritations et pathologies laryngées consécutives au reflux gastro-œsophagien, les nodules, les kystes, les varicosités, l'asymétrie et la faiblesse.

Les études qui tentent de répondre à cet objectif utilisent souvent sur un nombre limité de sujets, recourent à l'observation par le fibroscope souple, et ne s'appuient pas sur des modèles expérimentaux déterminés.

2.1.4. La fiabilité du diagnostic

Dejonckere et al (1998)^[64] ont réalisé une étude portant sur la fiabilité de l'évaluation vidéolaryngostroboscopique.

Le but de leur travail est d'optimiser un protocole d'enregistrement vidéostroboscopique laryngé pathologique. Le protocole utilisé par Dejonckere et al a été établi par le B.S.G.V.D. [annexe 2]. Il présente 30 enregistrements

vidéostroboscopiques de diverses pathologies laryngées à 7 O.R.L. expérimentés. Le degré de concordance peut être considéré comme élevé pour les paramètres suivants : l'amplitude (0.90), l'ondulation de la muqueuse (0.75), le type d'insuffisance d'occlusion glottique (0.70) et d'asymétrie (0.92). Il est moyen pour les paramètres de régularité (0.49), de consistance et de comportement vibratoire de la lésion (0.61). Il est faible pour le degré d'importance de l'insuffisance d'occlusion glottique et le degré d'asymétrie (0.27). Les évaluateurs expérimentés dans l'administration du protocole, présentent entre eux un plus haut degré de concordance que ceux qui ne disposent pas de cette expérience et ce, particulièrement pour les paramètres de fermeture glottique (0.54 vs 0.18) et d'ondulation de la muqueuse (0.75 vs 0.56). Dejonckere et al^[64] notent également que les variables catégorielles ont un score de reproductibilité inférieur aux variables quantitatives. Le paramètre de régularité montre une corrélation inter-juges faible mais intra-juges élevé.

D'autres protocoles sont proposés et des études sont également réalisées pour tester la fiabilité des évaluateurs. Ramos et al (1993)^[244] rapportent une fiabilité inter-juges qui s'étend de 0.20 à 0.85 et Poburka (1999)^[236] une fiabilité 0.21 à 0.90. Teitler (1995)^[318], quant à eux, montrent une fiabilité intra-juges de 0.31 à 0.97, et une fiabilité inter-juges de 0.75 à 0.98.

Hirano et al (1993)^[127] identifient un certain nombre de facteurs qui peuvent influencer le processus d'obtention et d'interprétation des images stroboscopiques, par exemple :

- la connaissance des observateurs concernant le lien entre la vibration des cordes vocales et le son produit,
- la connaissance de l'anatomophysiologie normale,
- l'habileté dans l'interprétation des images enregistrées.

Bless et al (1987)^[24] indiquent que bien que les compétences s'acquièrent avec un entraînement de base relativement court, quelques aspects de la stroboscopie rendent son utilisation délicate. Néanmoins, Poburka^[236] pense que la fiabilité pourrait être améliorée si la forme de l'évaluation était modifiée.

Poburka^[236] propose une étude basée sur un nouveau protocole baptisé S.E.R.F. (Stroboscopy Evaluation Rating Form) [annexe 3]. Il espère améliorer la

fiabilité des observateurs et compare ses résultats avec ceux obtenus par la forme S.A.V. (Stroboscopic Assessment of Voice) [annexe 4].

Les résultats obtenus avec S.A.V. ou avec S.E.R.F. sont identiques, 86 % de concordance pour l'un et 83 % pour l'autre. Le S.E.R.F. apporte des améliorations sur les paramètres suivants : l'activité supra-glottique, le bord libre de la corde vocale, la phase de fermeture, la phase de symétrie, la régularité, et la fermeture glottique. L'auteur présente 42 échantillons vidéostroboscopiques à 3 O.R.L. expérimentés. Les exemples sont échantillonnés d'une phonation normale à une dysfonction phonatoire sévère. Deux modes de concordance sont évalués : le pourcentage de concordance exact qui correspond à une évaluation identique des 3 juges et le pourcentage de consensus qui correspond à l'évaluation identique de 2 juges. Le pourcentage de concordance est modéré. Celui de consensus est bon. Les paramètres ayant un degré de concordance élevé sont le niveau vertical et la portion non vibrante (71%, 76%). L'auteur explique ce haut degré de fiabilité par le fait que peu d'échantillons vocaux ont une partie non vibrante, et peu aussi présentaient une phase verticale anormale. L'accord des juges cependant est plus faible pour les paramètres suivants : l'onde de la muqueuse (6%), la phase de fermeture (24%), la phase de régularité (12%). Les modifications apportées par le S.E.R.F. ne résolvent pas suffisamment les problèmes d'évaluation de ces paramètres. Néanmoins l'auteur souligne que la nouvelle forme apporte un certain nombre d'avantages : une évaluation plus rapide, moins d'hésitations, moins de confusion dans l'orientation des formes. En moyenne, le temps d'évaluation pour chaque cas est de 4 à 5 minutes.

Comparer le degré de concordance des évaluateurs aux différents protocoles est difficile, les paramètres à évaluer n'étant pas présentés de la même manière. Cependant, sur un paramètre tel que l'ondulation de la muqueuse, le degré de concordance rapporté par Dejonckere et al^[64] et par Poburka^[236], est radicalement opposé, l'un obtient un degré de concordance élevé dans le protocole du B.S.G.V.D., l'autre un degré faible dans le S.A.V. et le S.E.R.F..

La vidéostroboscopie dépend donc de l'œil averti de l'O.R.L., et reste l'outil de prédilection pour l'observation du larynx. Le chapitre qui suit introduit les échelles d'auto-évaluation qui ont pour but de cerner au mieux les plaintes du patient.

Chapitre 3

L'évaluation subjective

3.1. LES ÉCHELLES D'AUTO-ÉVALUATION.

Ces échelles permettent au thérapeute d'être informé de la manière dont le patient perçoit son émission vocale, du pourcentage d'invalidité qu'il s'attribue dans différents lieux et circonstances, et dans quel sens il souhaiterait que son émission soit modifiée.

3.1.1. L'échelle bipolaire d'auto-estimation vocale

Cette échelle, mise au point par Dejonckere et al (1985)^[58] consiste à demander au patient ou au sujet de porter un jugement sur la qualité de son émission vocale. Le sujet doit évaluer sur une échelle à 7 grades sa voix actuelle et celle qu'il aimerait obtenir.

Paramètres	1	2	3	4	5	6	7	Paramètres
Fatiguée								Reposée
Pénible								Facile
Sourde								Sonore
Inexpressive								Expressive
Contractée								Décontractée
Fabriquée								Naturelle
Désagréable								Agréable
Rauque								Pure
Engorgée								Posée
Faible								Puissante
Féminine								Masculine

Légende : 1 : très mauvaise ; 2 : mauvaise ; 3 : plutôt mauvaise ; 4 : impossibilité de répondre ; 5 : plutôt bonne ; 6 : bonne ; 7 : très bonne.

L'analyse peut être réalisée lors de l'examen avec le thérapeute. La voix peut être évaluée dans différentes circonstances : en voix conversationnelle, à forte ou faible intensité, en lecture, en comptage, en voix d'appel, sur une voyelle tenue avec étude de l'attaque, de la durée, de la fin d'émission, en voix chantée.

Le but est d'objectiver l'examen en chiffrant les paramètres.

3.1.2. L'index de handicap vocal ou V.H.I. (Voice Handicap Index).

Cette échelle d'auto-évaluation a été mise au point par Jacobson et al (1997)^[152]. Elle est née de la nécessité de quantifier l'aspect "invalidant" qu'un trouble vocal peut provoquer sur la qualité de vie, et de mesurer les conséquences psychosociales des troubles vocaux. L'objectif des auteurs est également de développer une échelle applicable à une grande variété de troubles vocaux. La version initiale comprenait 85 items, réduite dans une forme ultérieure à 30 items. Ces items sont groupés en trois catégories liées aux aspects fonctionnel, émotionnel et physique. L'échelle de cotation est établie sur 5 points (0-4)¹⁵. La dernière version a été éprouvée au test de reproductibilité sur 63 patients dont l'âge moyen est de 49 ans. Elle s'avère particulièrement robuste avec un degré de corrélation élevé ($\alpha=0.95$). Les items fonctionnels sont évalués avec une corrélation de $r=0.84$, les items émotionnels $r=0.92$, et les items physiques $r=0.86$.

Pour compléter leurs observations, les auteurs ont corrélé les résultats au degré de sévérité global du désordre vocal. Les patients évaluaient la sévérité de leur trouble vocal sur une échelle de 0 à 3 (0 correspondant à la voix normale, 1 à la voix légèrement diminuée, 2 à la voix moyennement diminuée, 3 à la voix sévèrement diminuée). La corrélation est modérée ($r=0.60$).

Le protocole de passation de l'échelle se trouve en annexe 5.

Benninger et al (1998)^[19] corrélient le V.H.I. au SF-36. Le SF-36 correspond à une mesure de la "qualité générale de la vie" présentée sous forme d'un questionnaire. Les auteurs mesurent 8 domaines de la santé communément modifiés par le trouble et son traitement : la fonction physique (les limites physiques dans l'exercice de la marche, de la baignade, des sports énergiques (par exemple le squash...), le rôle de la fonction physique (les difficultés rencontrées au travail, dans des tâches quotidiennes), la douleur corporelle (son intensité et les limites qu'elle provoque), la santé générale, la vitalité (le niveau d'énergie), le fonctionnement social (les interférences de l'état de santé sur des activités sociales), la santé mentale et la

¹⁵ Pour la cotation du V.H.I.: 0 correspond à jamais, 1 à presque jamais, 2 à parfois, 3 à presque toujours, 4 à toujours.

fonction émotionnelle (les difficultés rencontrées au quotidien résultant de problèmes émotionnels). La population étudiée comprend 260 adultes âgés de 18 à 90 ans, composée à 64.5% de femmes et appariée en fonction de leur trouble vocal. Ils comparent également ses résultats avec des résultats antérieurs rapportés au SF-36 sur des sujets normaux et des patients souffrant de sinusite chronique, d'angine, de sciatique, de défaillance cardiaque et de trouble pulmonaire chronique. Les troubles vocaux ont bien un impact significatif sur la qualité de vie des patients ; la fonction sociale et l'habileté à produire des tâches au quotidien sont plus affectées. La corrélation la plus haute se situe entre la fonction sociale du SF-36 et le score global du V.H.I. ($p<0.001$). Une corrélation significative est obtenue entre le domaine de la santé mentale ($p<0.01$), de la santé en général ($p<0.01$), celui du rôle de la fonction émotionnelle ($p<0.017$) et le score globale du V.H.I.. Les patients dysphoniques ont un niveau de fonctionnement physique inférieur à celui des patients souffrant de sinusites chroniques, ($p<0.01$), et subissent un plus grand handicap. La dysphonie entraîne également des niveaux de fonction sociale plus bas que l'angine ($p<0.01$), la sciatique ($p<0.01$) et un score de santé mentale plus bas que l'angine ($p<0.01$). Les auteurs trouvent des corrélations entre le SF-36 et le V.H.I. dans les domaines de la fonction sociale, de la santé mentale, et du rôle de la fonction émotionnelle.

Le V.H.I. a l'avantage de quantifier l'impact d'une variété de troubles vocaux sur la qualité de la vie. Sa fiabilité est élevée (0.95) ainsi que ses liens avec des échelles psychologiques. Son utilisation se répand et satisfait les cliniciens des laboratoires de recherche. Il enrichit l'anamnèse du patient qui canalise rapidement les difficultés rencontrées dans sa vie quotidienne et s'adapte bien aux réalités de la consultation phoniatrique.

3.2. LES ÉCHELLES PERCEPTUELLES.

L'évaluation la plus ancienne et probablement la plus pratiquée est le jugement perceptuel. Elle trouve ses origines dans la description propre et la terminologie des professeurs de chant, des pédagogues de la voix. Ils parlent de ce qu'ils entendent. Cette technique est très couramment retrouvée dans la littérature (in De Bodt et al (1996)^[51]). Les controverses concernent la terminologie, la méthodologie et la fiabilité. Les chercheurs soulignent son importance mais se rendent compte de ses limites. On reconnaît son utilité dans l'appréciation subjective de la thérapie vocale.

De Bodt en citant Perkins (1971) souligne l'incohérence d'une terminologie abondante tentant de décrire la voix. Cet auteur, dans une comparaison de 9 études ne relève pas moins de 27 termes pour décrire la voix des locuteurs. Les termes communs aux différentes études sont la raucité et la nasalité. Ce qui explique que plusieurs centres ont créé leur propre jugement perceptuel, qui ne répond qu'à un consensus interne à leur laboratoire.

3.2.1. Les profils décrits dans la littérature

La nécessité d'une terminologie adéquate et d'une méthode descriptive d'évaluation de la qualité vocale conduit au développement d'un protocole standardisé. Hammarberg et al (1980)^[108], (1984)^[109], 1986^[110]; Wilson et al (1973)^[345], Laver (1980)^[177], (1986)^[178] et la Société Japonaise de Logopédie et de Phoniatrie ^[122], parmi d'autres ont élaboré de tels protocoles. Cependant aucun système n'a encore été universellement reconnu ; l'échelle G.R.B.A.S. (Grade, (grade), Rough (raucité), Breath (souffle), Astheny (sensation de faiblesse), Strain (sensation de forçage) s'imposerait.

3.2.1.1. Le Buffalo Voice Profile de Wilson (B.V.P.) (1987) ^[346]

Le but du B.V.P. est de fournir des critères et des profils d'après lesquels sont évalués différents paramètres vocaux. Le B.V.P. comporte 10 profils :

1. des profils de dépistages vocaux (employés pour dépister les catégories vocales dans une large population d'enfants),
2. un profil de groupes comportementaux (observation et évaluation des enfants sur des séries simples),
3. un profil vocal (instructions générales pour la thérapie vocale),
4. un profil d'abus vocaux (basé sur une liste de 11 types communs d'abus vocaux)
5. un profil du locuteur anxieux (rigidité, inhibition, dysfluence (dysfluency), et agitation),
6. un profil de résonance (12 items pour juger les problèmes de résonance),
7. un profil vocal d'audition altérée,
8. un profil d'enregistrements vocaux (réalisé au début, en cours et à la fin du programme de réhabilitation),
9. un profil de diagnostic vocal (tous les résultats, incluant les mesures de la voix, sont rapportés sur un seul profil).
10. un profil de progrès en thérapie vocale

Ces profils sont composés de 12 paramètres évalués par les moyennes d'une échelle à 5 points.

Les paramètres sont les suivants :

1. le son laryngé (souffle, raucité, dureté),
2. la hauteur tonale (trop haut, trop bas),
3. l'intensité (trop faible, trop fort),
4. la nasalité (hypernasalité, hyponasalité),
5. la résonance orale (raucité),
6. le souffle (quantité),
7. les muscles (hypertendus, hypotendus),
8. les abus vocaux (quantité et degré),

9. la vitesse (trop rapide, trop lente),
10. l'anxiété dans le discours,
11. l'intelligibilité du discours,
12. l'estimation vocale globale.

L'échelle d'évaluation de la sévérité s'échelonne de 1 à 5 comme suit :

1. voix normale,
2. sévérité moyenne,
3. sévérité modérée,
4. problème sévère,
5. problème très sévère.

Trois éléments essentiels ressortent de l'évaluation vocale : la sévérité du problème, les troubles de la communication et l'indication posée pour l'intervention. Les échantillons langagiers contiennent du langage spontané, une lecture à haute voix, l'émission de phonèmes isolés, et du comptage. L'inconvénient est de ne pas avoir d'exemples enregistrés des paramètres qu'ils identifient. La fiabilité intra et inter-juges ne peut donc pas être garantie.

Wilson et al^[in 51] recommandent l'échelle d'intervalles à apparence égale (E.A.I.) pour évaluer la qualité vocale. Ces échelles contiennent principalement de 3 à 7 points. Un désavantage de l'échelle est "the end-effect", l'évaluateur ayant tendance à éliminer les cotations situées aux extrémités de l'échelle. Il en réduit donc le nombre de degrés. De plus, L'évaluateur ne dispose d'aucune définition des points de l'échelle.

3.2.1.2. Le schéma de Hammarberg (1992)^[111].

Selon Hammarberg, la qualité vocale est conditionnée par un langage culturel commun et spécifique. L'auteur a essayé de déterminer un nombre limité de paramètres pour évaluer la qualité vocale et les a illustrés par des enregistrements prototypiques. Ils sont au nombre de 12 soit :

1. **l'aphonie, l'aphonie intermittente** : la voix est sans cesse interrompue par un manque de sonorisation, il y a des moments de chuchotements ou de pertes vocales,
2. **le souffle** : la fuite audible de l'air à travers la glotte est due à une fermeture glottique insuffisante,
3. **l'hyperfonctionnement/la tension** : la voix est "tendue", comme si les cordes vocales étaient comprimées durant la phonation,
4. **l'hypofonctionnement/le relâchement** : il est opposé à l'hyperfonctionnement. Il s'agit d'une insuffisance de tension cordale, résultant d'une voix relâchée et fragile,
5. **"vocal fry"/"creaky"** : il s'agit de séries rapides de claquements, de vibrations périodiques à fréquence grave,
6. **la raucité** : elle est caractérisée par un bruit a-périodique dont la hauteur tonale est basse. Elle est vraisemblablement liée à la vibration irrégulière des cordes vocales,
7. **la raucité avec une hauteur tonale haute** : le bruit est a-périodique. La hauteur est élevée et vraisemblablement liée à une vibration irrégulière des cordes vocales,
8. **les cassures vocales** : il s'agit de cassures fréquentielles intermittentes,
9. **la diplophonie** : deux hauteurs tonales différentes sont perçues simultanément.
10. **le registre** : les modes de phonation sont le modal et le falsetto,
11. **la hauteur** : elle est le corrélât auditif principal de la fréquence fondamentale,
12. **l'intensité** : elle est le corrélât auditif principal du niveau de pression sonore.

Hammarberg a aussi établi des corrélations entre les procédures d'évaluation standardisées et des évaluations acoustiques (évaluations spectrales). A partir des résultats de l'analyse factorielle et des corrélations perceptuelles et acoustiques, 12 paramètres minimum adéquats semblent être requis pour la description des voix pathologiques suédoises. Ces paramètres sont présentés comme cliniquement pertinents et recommandés dans la pratique de l'évaluation quotidienne du trouble vocal. L'auteur souligne que l'évaluation est fiable lorsqu'elle est pratiquée par des auditeurs bien entraînés et que les procédures d'évaluation sont standardisées. L'efficacité des auditeurs est augmentée lorsque les relations acoustico- perceptuelles sont identifiées. En ce qui concerne les tâches, Hammarberg choisit du langage spontané et la voyelle tenue. Les modifications dans le langage spontané telles que le

début et la fin de l'émission sonore, les cassures vocales sont décisives pour la qualité vocale, ce que le test de la voyelle ne permet pas. L'auteur emploie un texte standard de 40 secondes. Laver et al (1986)^[178] comme Hammarberg^[111] préfèrent nuancer davantage la description de la qualité vocale par des termes multiples plutôt que par des termes simples. Laver et al préfèrent parler de voix soufflée hyperfonctionnelle accompagnée de moments d'aphonie intermittents plutôt que d'utiliser le terme brut "soufflé". L'avantage de ces expressions est que l'évaluation de la qualité vocale ne se réduit pas à un seul élément.

En dépit des efforts de l'auteur pour établir des corrélations acoustico-perceptuelles, cette méthode d'évaluation n'a pas été adoptée.

3.2.1.3. Le profil vocal de Laver^[178]

Laver propose un outil perceptuel issu de la phonétique et appelé "protocole d'analyse" du profil vocal. Il définit la qualité vocale au sens large comme *"la coloration auditive caractéristique de la voix à l'audition et non pas comme dépendant de la seule activité laryngée"*. Le système décrit aussi bien des paramètres laryngés que supra-laryngés (labial, mandibulaire, lingual ou vélo-pharyngé). Chaque caractère est rapporté à une "ligne de base" neutre et non pas à un concept fictif de normalité. Les lignes de base neutres ont des corrélats physiologiques et acoustiques bien circonscrits. Nous présentons le protocole en anglais (tableau 17) vu le nombre d'expressions difficiles à traduire. Il est divisé en trois parties : les caractéristiques vocales, les caractéristiques prosodiques et l'organisation temporelle. Le jugement est basé sur l'observation auditive et visuelle. L'écartement du point neutre peut être estimé selon 6 degrés. Les types de voix sont classés comme rauques, chuchotés, soufflés, grinçants, voix de tête, et voix usuelle. L'impression générale de la qualité vocale est envisagée comme la résultante de plusieurs composants potentiels indépendants entre eux. Aucun score pour cette impression générale n'est prévu.

Tableau 17 : "Voice Profile Analysis Scheme" de Laver^[179].

1. Vocal quality features		2. Prosodic features	
Vocal tract			
a. Labial	Lip rounding / protrusion	a. Pitch	High mean
	Lip spreading		Low mean
	Labiodentalization		Wide range
	Extensive range		Narrow range
	Minimized range		High variability
b. Mandibular	Close jaw		Low variability
	Open jaw	b. Loudness	High mean
	Protruded jaw		Low mean
	Extensive range		Wide range
	Minimized range		Narrow range
c. Lingual top/blade	Advanced		High variability
	Retracted		Low variability
d. Lingual body	Fronted body		
	Backed body	3. Comments	
	Raised body		
	Lowered body	a. Breath support	
	Extensive range	b. Continuity	Interrupted
	Minimized range	c. Rate	Fast
e. Velopharyngeal	Nasal		Slow
	Audible nasal escape	d. Rhythmically	
	Denasal	e. Other	
f. Pharyngeal	Constriction		
g. Larynx position	Raised		
	Lowered		
Phonation type	Harshness		
	Whisper(y)		
	Creak(y)		
	Falsetto		
	Modal voice		
Tension			
h. Supralaryngeal	Tense		
	Lax		
i. Laryngeal	Tense		
	Lax		

3.2.1.4. L'échelle G.R.B.A.S de Hirano (1981)^[122]

Ce sont les travaux d'Isshiki et al (1966)^[147], (1969)^[148], (1970)^[149] sur les différentes méthodes sémantiques de Osgood (1957)^[224] pour quantifier les impressions auditives de raucité qui donnent naissance à l'échelle G.R.B.A.S.. Il définit 17 adjectifs et leur contraire à partir de 260 termes sélectionnés dans la littérature. Seize échantillons de voix rauque sont analysés par un jury de professionnels et de non-professionnels. Des analyses de données se dégagent 4 facteurs : la raucité (**R**), le souffle (**B**), la sensation de faiblesse (**A**) et la proximité du normal (**N**). Il apparaît clairement qu'aucun de ces facteurs ne suffit à lui seul. Les résultats de l'analyse factorielle met en évidence pour les deux juges ces quatre mêmes facteurs. La classification (**R**), (**B**), (**A**) semblait plus facile dans les cas de raucité sévère. Le facteur (**N**) lié avec la raucité change d'appellation et devient le facteur (**D**) pour dysphonie (Isshiki et al (1969)^[148]).

Ces 16 échantillons vocaux sont envoyés dans plusieurs centres médicaux du Japon et évalués de 2 manières. La première méthode, plus facile, consiste à attribuer un degré à la raucité. La seconde qui use d'une échelle à 4 points (0 = normal, 1 = léger, 2 = moyen, 3 = extrême) pour tous les paramètres, exigeait une différenciation plus fine de la raucité. Dans le cadre d'une étude sémantique différentielle, Hiroto^[in 122] applique une analyse spectrographique aux 16 échantillons. Le paramètre (**B**) est caractérisé par des bruits s'étendant des fréquences basses aux fréquences élevées, le paramètre (**R**), par une composante en harmoniques pauvres, et des variations d'amplitude et de fréquence fondamentale. Le paramètre (**A**) se rapporte à un affaiblissement de l'intensité. Isshiki et al (1969)^[148] montrent que le caractère (**R**) est caractérisé par une grande perturbation de la hauteur tonale, tandis que le facteur (**B**) se distingue par une composante de bruit et une composante d'harmoniques réduites ou négligeables.

Le Comité des tests de la fonction phonatoire de la Société Japonaise de Logopédie et de Phoniatrie a davantage développé le travail d'Isshiki et al comme l'échelle G.R.B.A.S. pour décrire les voix anormales.

Dans la description de la qualité vocale, le système de l'échelle G.R.B.A.S. contient 5 paramètres bien définis :

1. **Le grade (G)** : correspond au degré de raucité ou d'anormalité vocale.
2. **La raucité (R)** : donne l'impression d'irrégularités de vibration des cordes vocales. Cet effet correspond à la fluctuation irrégulière de la fréquence fondamentale et/ou de l'amplitude de la source sonore glottique.
3. **Le souffle (B)** : rend compte de l'étendue de la fuite d'air à travers la glotte, liée à la turbulence.
4. **La sensation de faiblesse (A)** : traduit un manque de puissance dans la voix, lié à une intensité faible ou au manque d'harmoniques aigus.
5. **La sensation de forçage (S)** : traduit un état hyperfonctionnel de la phonation. Cet état est lié à une fréquence fondamentale anormalement élevée et à du bruit dans les fréquences aiguës.

Chaque paramètre est coté sur une échelle à 4 points. 0 correspond à l'état normal, sans raucité, 1 à une raucité légère, 2 à une raucité moyenne et 3 à une raucité sévère.

De tous les systèmes d'évaluation, cette échelle est la plus compacte. Elle est conçue pour un emploi journalier par tous les membres de l'équipe médicale. La Société Japonaise a réalisé des enregistrements standards de voix typiques illustrant chaque paramètre de l'échelle G.R.B.A.S. Ils sont accompagnés de commentaires sur la fonction laryngée avec une attention particulière pour les paramètres supra-laryngés. Une fois l'auditeur entraîné, la pratique de l'échelle est rapide et facile.

L'échelle ne fournit aucune information concernant la hauteur, l'intensité ou la résonance comme le font la plupart. Cette lacune est corrigée par Hirano (1990)^[126] qui complète l'évaluation par des paramètres de hauteur (haute, normale, basse), d'intensité (forte, normale, faible) et ajoute celle de qualités spécifiques telles que les cassures vocales, l'aphonie, l'hyperfonctionnalité, l'hypofonctionnalité, les voix spasmodiques, le tremblement vocal (tremor). La hauteur et l'intensité sont mesurées par des instruments objectifs. Dejonckere (1995)^[62], entraîné à l'emploi de l'échelle, introduit un sixième facteur : l'instabilité de la raucité (I).

Au vu des quatre systèmes exposés, De Bodt et al (1996)^[51] recommandent l'emploi de l'échelle G.R.B.A.S. et l'ajout de mesures vocales objectives. Ils soulignent sa facilité d'emploi dans un environnement clinique journalier et l'avantage qu'elle a de

fournir une description de la qualité vocale relativement cohérente. L'échelle peut également être employée comme telle par chaque professionnel quelles que soient la profession ou l'expérience des utilisateurs.

3.2.2. La fiabilité de l'échelle.

Kreiman et al (1990)^[169] répartissent les problèmes méthodologiques en 3 catégories : les paramètres étudiés, la qualification du jury (en terme d'expérience, d'entraînement) et le contenu phonétique de l'échantillon.

La hauteur, l'intensité et le rythme, entre autres, peuvent influencer le jugement des auditeurs et leur fiabilité.

3.2.2.1. La qualification du jury.

Anders et al (1988)^[4] qui étudient l'effet de l'entraînement et du bagage culturel sur la perception de la raucité, ne trouvent pas de facteurs prédominants. Hammarberg^[108, 110] a observé un bon niveau de fiabilité pour un groupe d'auditeurs expérimentés, ce que confirme Gelfer (1988)^[87] pour qui des auditeurs expérimentés jugent de façon plus cohérente que des évaluateurs non entraînés.

Kreiman et al (1990)^[169] montrent que les auditeurs, selon qu'ils sont naïfs ou expérimentés emploient des stratégies perceptuelles différentes. Pour eux, les juges experts atteignent un pourcentage d'accord moindre que les juges naïfs en ce qui concerne l'importance relative des différents aspects de la qualité vocale. Ils suggèrent donc la plus grande prudence dans l'emploi de données moyennes obtenues sur un échantillon dispersé de cliniciens. En outre, kreiman et al recommandent l'emploi de stimuli d'ancrage pour contrôler la variabilité de l'évaluation vocale. Bassich et al (1986)^[18] rapportent que 8 heures d'entraînement sont requises pour atteindre 80% de fiabilité inter-juges en employant un système d'évaluation à 13 dimensions. De Bodt et al (1997)^[53] rapportent à ce sujet que l'expérience et le bagage professionnel n'influencent pas l'évaluation des paramètres s'ils comparent les

scores de la médiane (niveau de significativité pour le degré d'expérience : **(G)**, $p=0.18$; **(R)**, $p=0.7$; **(B)**, $p=0.32$; **(A)**, $p=0.32$; **(S)**, $p=0.6$ et le niveau de significativité pour le bagage professionnel : **(G)**, $p=0.36$; **(R)**, $p=0.11$; **(B)**, $p=0.18$; **(A)**, $p=0.32$; **(S)**, $p=0.11$). Néanmoins, s'ils observent les résultats basés sur les moyennes, les scores des auditeurs expérimentés sont plus élevés que les scores des auditeurs inexpérimentés ($k=0.70$ vs 0.50). Les ORL expérimentés fournissent la meilleure moyenne globale. Ils ajoutent que les auditeurs expérimentés ont tendance à surestimer le degré de dysphonie tandis que les auditeurs non expérimentés la sous-estiment. Le consensus général est que l'expérience est un facteur-clé pour l'obtention d'une analyse perceptuelle fiable (Hammarberg, De Bodt, Anders, Askenfeld, Kreiman^{[in [256]]}).

3.2.2.2. Les paramètres de l'échelle.

Abe et al (1986)^[1] jugent que la reproductibilité de l'échelle G.R.B.A.S. diffère selon les auditeurs et les paramètres de l'échelle.

Les travaux de Dejonckere et al (1993)^[61] montrent que dans une liste de score perceptuel de 15 items, l'échelle G.R.B.A.S. apparaît la plus intéressante. Les auteurs repèrent des combinaisons favorables de variances intra, inter-juges et des variances inter-vocales élevées. Le paramètre **(G)** ($p=0.7$) obtient la meilleure corrélation entre 2 juges. Les paramètres A et S offrent moins de cohérence.

De Bodt (1997)^[52] trouve une fiabilité inter-juges (ORL et logopèdes) modérée ($K=0.43$). Le meilleur accord est obtenu pour le paramètre **(G)** (0.39) et le plus mauvais pour le paramètre **(S)** (0.11).

Leinonen et al (1997)^[181] cherchent à mettre en évidence les caractéristiques communes aux évaluations des cliniciens expérimentés et à déterminer les dimensions principales de la variation acoustique et perceptuelle. Six logopèdes expérimentés jugent 94 échantillons de voix masculines et 124 de voix féminines sur une échelle visuelle analogique de 5 cm. Ils évaluent le degré de pathologie des paramètres suivants : la raucité **(R)**, le souffle **(B)**, la sensation de forçage **(S)**, et sensation de faiblesse **(A)**. La concordance entre les juges est la plus élevée pour le

degré de pathologie (0.734), le souffle (0.698) et la raucité (0.646). La pathologie est d'ailleurs principalement déterminée par le souffle et la raucité.

A partir d'une revue de la littérature explorant 57 études de 1951 à 1990, Piccirillo et al (1998)^[231] rapportent une haute fiabilité inter-juges pour chaque paramètre de l'échelle G.R.B.A.S. ((**G**) : 0.96 ; (**R**) : 0.95 ; (**B**) : 0.96 ; (**A**) : 0.92 ; (**S**) : 0.93).

Kreiman et al (1990)^[169] démontrent la supériorité de l'échelle G.R.B.A.S. par rapport à d'autres méthodes d'analyse perceptuelle pour lesquelles le niveau de reproductivité et de fiabilité variait de 18 à 100%. C'est le paramètre (**G**) de la dysphonie qui semble récolter un niveau de corrélation le plus important.

Olson et al (1998)^[220] évaluent les résultats acoustiques et perceptuels d'une réinnervation par anastomose de l'anse cervicale sur le nerf récurrent. Cinq juges apprécient les paramètres G.R.B.A.S. sur une échelle de 12mm. La fiabilité intra-juges est excellente. Les patients montrent une amélioration sur les paramètres suivants : le Grade (**G**), le Souffle (**B**) et la sensation de forçage (**A**) ($p < 0.05$). Olson et al soulignent que le grade est la mesure la plus simple et la plus adéquate pour évaluer la qualité vocale, tout en étant aussi la mesure la mieux corrélée avec le degré de satisfaction du patient.

Millet et al (1998)^[201] rapportent des valeurs de fiabilité de 2 juges sur l'échelle G.R.B.A.S. concernant l'évaluation de 65 sujets dysphoniques. La corrélation intra-juges est plus élevée que la corrélation inter-juges. Le grade est plus fiable que le souffle et la raucité ((**G**) : 0.87 ; (**R**) : 0.70 ; (**B**) : 0.69).

Carding et al (1999)^[35] confirment l'efficacité d'une thérapie vocale chez 45 patients souffrant d'une dysphonie dysfonctionnelle. Quatre juges ont évalué le degré de sévérité de la dysphonie en se basant sur un seul paramètre de sévérité globale à partir de 90 échantillons vocaux. La fiabilité intra-juges s'étend de 0.94 à 0.99 indiquant un accord presque parfait. La cohérence inter-juges est de 0.89 ($p < 0.05$).

3.2.2.3. L'échantillon.

Le matériau phonétique soulève des questions dans la mesure où les mesures objectives sont réalisées sur une voyelle tenue tandis que les mesures subjectives le sont sur un fragment de parole (Hammarberg et al (1980)^[108]). A ce propos, les points

de vue divergent en fonction des auteurs. Les voyelles tenues sont limitées à leur partie stationnaire ce qui en fait un matériau appauvrissant pour l'analyse vocale. Il reste cependant l'apanage de tous (Kempster et al (1991)^[161].

Rabinov et al (1995)^[241]) Les arguments en faveur des voyelles sont plus pragmatiques et concernent plutôt les mesures objectives. Les arguments qui défendent le fragment de parole articulée sont plus théoriques et concernent l'analyse subjective.

Le discours spontané est souvent considéré comme le matériau le plus approprié étant donné qu'il contient un phénomène transitoire permettant l'analyse des dynamiques vocales en condition naturelle (Hammarberg^[108]). La portion stable d'une voyelle isolée (le plus souvent [a]) offre des avantages pratiques pour le chercheur (Kreiman, Bassich, Childers, Wolfe,^[in 256]). Elle n'est pas affectée par le facteur de co-articulation. Elle est utile dans l'analyse vocale objective pour l'obtention des mesures acoustiques et aérodynamiques (Giovanni 1994/1995^[93]). Cependant, Bassich^[256] souligne le caractère artificiel de ce matériau qui peut conduire à une sous-estimation des changements pathologiques. Sakata et al (1995)^[263] montrent une différence dans l'évaluation perceptuelle d'une voyelle soutenue et du langage courant produits par différentes voix. Le Grade (**G**) est obtenu une valeur plus élevée dans l'échelle pour la langage courant. La sensation de faiblesse (**A**) et la sensation de forçage (**S**) sont évaluées plus haut dans la voyelle soutenue.

de Krom et al (1994)^[54] demandent à 6 juges d'évaluer le souffle et la raucité de 78 sujets dysphoniques. Les échantillons évalués perceptuellement comprennent la portion stable et la portion complète des voyelles (attaque et finale de l'émission comprises) ainsi que du langage continu. La cohérence est plus élevée entre le langage continu et les voyelles complètes qu'entre le langage continu et les voyelles stables. de Krom et al ne notent aucun effet significatif du choix de l'échantillon de parole, sur la concordance intra ou inter-juges dans l'évaluation du Grade, de la Raucité ou du Souffle. Revis et al (1999)^[256] corroborent ces observations. Ils demandent à 7 juges d'évaluer 60 sujets dysphoniques et 20 sujets contrôles au moyen de l'échelle G.R.B.A.S.. Les 3 tâches sont identiques à celles proposées par de Krom et al^[54]. Ils observent que les juges sont moins sévères lorsqu'ils jugent une voyelle soutenue stable qu'un échantillon de langage spontané, mais que les jugements sur la voyelle soutenue complète sont similaires aux jugements portés sur le langage spontané. A l'analyse perceptuelle, les juges portant leur évaluation sur la

portion initiale et stable d'une voyelle soutenue ne sous-estiment pas la dysphonie mais Revis et al soulignent que la production dans des conditions artificielles peut être un choix qui prête à confusion. Ils considèrent donc que le test de la voyelle tenue est sans doute pertinent pour une analyse vocale mais inadéquat pour une analyse perceptuelle.

de Krom et al (1995)^[55] ont demandé à un jury d'analyser la qualité vocale de sujets normaux et dysphoniques sur base de 4 échantillons vocaux :

un fragment de parole, les 200 premières millisecondes (attaque) d'une voyelle, les 200 millisecondes de la partie stable de cette voyelle et la voyelle entière attaque incluse.

Ils rapportent que les fragments comprenant la voyelle dans sa totalité sont plus interprétables que la partie stable. Pour les sujets dysphoniques moyens, ce sont les parties instables du [a] (attaque et finale ou encore les transitoires d'entrée et de sortie) qui sont atteintes en premier.

Giovanni et al (1996)^[93] insistent sur le fait que toutes les étapes de l'évaluation perceptuelle doivent être analysées et contrôlées pour permettre des comparaisons valables. Les biais peuvent être l'expérience, la sensibilité de l'auditeur, le choix du matériau à analyser et l'échelle choisie. En outre, les stratégies d'écoute des auditeurs doivent être approfondies pour améliorer leur fiabilité.

L'échelle G.R.B.A.S. permet d'évaluer la qualité vocale de sujets présentant pathologies diverses. Cependant, pour les troubles vocaux du chanteur, elle s'avère peu efficace. En effet, la voix parlée de certains patients dysodiques est souvent évaluée comme étant "normale", alors qu'en voix chantée le clinicien pourra percevoir du souffle, de la raucité ou encore de l'irrégularité vibratoire. Au point suivant, nous présentons donc une échelle perceptuelle qui est le fruit de notre travail personnel et qui a pour but d'étudier le vibrato.

3.2.3. Étude 1 : Une échelle perceptuelle pour l'étude du vibrato.

Nous commençons par une brève revue de la littérature concernant le vibrato pour ensuite entamer notre étude personnelle.

La suite de notre contribution se situe au chapitre 5. Cette première étude trouve sa place ici vu qu'il s'agit d'un sujet consacré aux évaluations perceptuelles du chapitre 3.

3.2.3.1. Introduction.

Les productions musicales instrumentales ont fait l'objet de plus d'évaluations perceptuelles que les productions musicales vocales. Cette différence est liée à la difficulté d'obtenir une évaluation fiable et valide des échantillons vocaux (Ekhlom et al (1998)^[71], Killian (1997)^[163], Wapnick et al ^[341]). Plusieurs études proposent des critères d'évaluation perceptuelle. Cependant, aucune échelle similaire à la G.R.B.A.S. n'a encore été proposée pour l'évaluation de la voix chantée. Les chanteurs forment pourtant une population non négligeable de la consultation en oto-rhino-laryngologie. Aussi, pour répondre à leur demande tant sur le plan diagnostic que rééducatif, le praticien a besoin de paramètres d'évaluation précis.

Le vibrato est une caractéristique de la voix chantée ainsi que du son produit par certains instruments de musique. Dejonckere et al (1995)^[63], dans l'introduction de leur ouvrage "Vibrato" soulignent l'importance du vibrato en citant les paroles du violoniste Carl Flesch : *"Le vibrato est considéré comme une question de vie ou de mort de l'artiste"*. D'autres types d'oscillations tels que le chevrottement et le trémolo qualifient les vibratos irréguliers. Cependant jusqu'il y a peu de temps, la terminologie était équivoque, une analyse de la littérature traduit cette confusion.

En 1936, Seashore^[282] définit un bon vibrato vocal comme *"une pulsation de hauteur tonale habituellement accompagnée de pulsations synchrones d'intensité et de timbre d'une étendue et d'une cadence telles que cela procure à la sonorité une agréable souplesse, un effet de tendresse et de richesse"*.

En 1938, Reid^[in 246] commence par décrire le trémolo instrumental obtenu par un violon comme une réitération rapide de l'archet sur une seule note, le vibrato instrumental comme une ondulation de la hauteur tonale par un mouvement de la main gauche. Le trémolo semble habituellement représenter l'ondulation, l'effet de fluctuation de la hauteur tonale et le vibrato, l'itération sur la même note.

Le "Grove's dictionary of music and musicians (5^{ème} édition)" crée la confusion. Le trémolo est décrit comme un exploit technique, le plus difficile à apprendre pour un chanteur^[in 246].

En 1967, Reid^[246] observe des confusions dans la terminologie. Ce qui est un trémolo pour l'un devient un vibrato pour l'autre et vice-versa.

En 1986, Miller^[200] parle du vibrato comme d'un terme assez vague utilisé pour décrire certaines sortes de fluctuations de hauteur tonale pouvant apparaître comme un son soutenu. Il en décrit les paramètres :

- Nombres d'ondulations par seconde [tableau 18]
- Fluctuation en hauteur tonale
- Variation d'intensité

L'auteur différencie trémolo, vibrato, chevrottement en fonction du nombre de pulsations par seconde et précise qu'une cadence inférieure à 4 pulsations par seconde est inacceptable pour l'oreille. Le son blanc est défini comme une voix chantée sans vibrato et utilisé dans beaucoup d'exercices et styles vocaux (le chant baroque notamment).

Tableau 18 : Nombre de cycles par seconde du vibrato calculé par différents auteurs.

Auteurs	Années	Cadences	Étendues
Miller ^[200]	1986	8 cycles/sec	
Rothman et al ^[260]	1987	5 à 6.7 cycles/sec	5 à 7
Hakes et al ^[107]	1990	5 à 7 cycles/sec	
Rothman et al ^[261]	1990	10 à 13 cycles/sec	3.7/sec
Cleveland et al ^[38]	1994	5.5 à 7 cycles/sec	1 à 2 demi-tons
Titze et al ^[326]	1994	5 cycles/sec	
Dejonckere ^[63]	1995	4.1 à 6.6 cycles/sec	0.38 à 3.26 demi-tons

En 1987, Rothman et al^[260] parlent de l'aspect culturel du vibrato, stipulant qu'il diffère d'une culture à l'autre et qu'il évolue dans une même société en fonction des valeurs esthétiques. Leurs commentaires sont subjectifs étant donné qu'aucune corrélation n'a été observée entre l'objectif et le subjectif.

En 1990, Rothman et al^[261] mentionnent que certains chanteurs peuvent produire un vibrato d'amplitude ou un vibrato de fréquence ou les deux.

En 1990, Hakes et al^[107] parlent d'un vibrato exagéré. C'est un vibrato où le chanteur permet à l'étendue de se développer au maximum. Ils précisent que les chanteurs peuvent tout changer, excepté leur cadence de vibrato. A ce sujet, Rothman note une définition du chevrottement tirée du New York Times¹⁶, traduite par le critique W. Crutchfield : *"le chevrottement est une caractéristique des voix âgées ou malades ; une vibration prononcée et lente au lieu d'une hauteur tonale maintenue fermement"*. Il spécifie également que *"le stress associé à une hauteur tonale ou à une intensité inadaptée à la hauteur de la voix ou à son développement technique, produit un chevrottement prématuré"*.

En 1991, D'Alessandro et al^[49] parlent du vibrato comme d'une fluctuation de la période de la fréquence fondamentale, de l'intensité, et du timbre du son. Ils observent que certains instruments montrent l'un ou l'autre aspect du vibrato :

l'orgue montre un vibrato d'amplitude engendré par le battement entre deux tuyaux légèrement désaccordés. Les instruments à cordes produisent un vibrato de fréquence. Le chant permet un vibrato plus complet mais où la vibration de fréquence semble dominer par rapport à celle d'intensité ou de timbre.

En 1994, Sundberg^[311] décrit le vibrato des sons chantés comme une ondulation régulière de la fréquence fondamentale caractérisée par deux variables : la cadence, soit le nombre de cycles ondulatoires par seconde et l'étendue, soit la grandeur du plus grand écart par rapport à la moyenne. Tout comme Hakes et al (1990)^[107], il ajoute que les chanteurs d'opéra semblent incapables de faire varier la cadence de leur vibrato.

En 1994, Cleveland^[38] décrit le vibrato comme un phénomène laryngé généré par le chanteur afin d'obtenir une modulation agréable et esthétique de la hauteur tonale et de l'intensité.

D'après Dejonckere et al (1995)^[63], le vibrato est produit par des modulations quasi périodiques de la fréquence fondamentale de la voix. Ils font l'hypothèse que

¹⁶ La date de parution est le 12 septembre 1986, pp 51-88.

l'agent principal du vibrato doit donc être un muscle ou un groupe de muscles qui régule le fondamental. Le muscle le plus important devrait être le crico-thyroïdien. Ce muscle montre des activités oscillatoires tant dans l'étude de Dejonckere que chez d'autres auteurs^[in 63]. Il en va de même pour les muscles crico-thyroïdiens latéraux, les thyro-aryténoïdiens et l'inter-aryténoïdien, soit les muscles glottiques intrinsèques.

Dejonckere remarque que les émissions de bons vibrati tendent à être associées à une taille adéquate de l'espace supra-glottique et hypo-pharyngé. Quant aux mécanismes dans le système nerveux central pour la production du vibrato, aucune information significative n'est apportée à ce jour.

Le vibrato est toujours associé à des oscillations synchrones de certaines parties du tractus vocal. Le voile du palais, la base de langue, l'épiglotte, et le mur pharyngé latéral proche du larynx oscillent souvent de manière synchrone avec le vibrato. Lorsque ces mouvements oscillatoires sont modérés et adéquats en étendue, ils semblent être utiles pour éviter la tension ou la rigidité de l'ajustement du tractus vocal. Lorsque les mouvements sont de petites amplitudes, l'effet est négligeable. Par contre, si les oscillations sont excessives, elles provoquent un vibrato défavorable. Dejonckere observe que le flux d'air oscille durant le vibrato. Ce flux ne serait probablement pas dû à la modulation de la pression d'air, mais aux modulations de la résistance glottique dépendant des actions de la musculature laryngée.

Les muscles respiratoires ne semblent pas exécuter d'oscillation durant le vibrato. Les muscles laryngés extrinsèques, incluant les muscles sterno-thyroïdiens et le sterno-hyoïdien, oscillent de manière synchrone avec le vibrato. Dejonckere^[in 63] ajoute que jusqu'à présent, les chercheurs ne sont pas capables de déterminer si ces muscles augmentent ou diminuent la fréquence fondamentale.

En 1995, Dejonckere et al^[63] synthétisent, dans leur ouvrage, l'apport de diverses études sur le vibrato. Dejonckere^[in 63] corrobore les résultats des études de la littérature par sa recherche sur 23 chanteurs. Il rapporte des cadences de vibrato qui s'étendent de 4.1 à 6.6 cycles par seconde. Les étendues de vibrato vont de 0.38 à 3.26 demi-tons. Aucune différence significative n'est observée en fonction des hauteurs tonales, des niveaux d'intensité et des types vocaux différents. Il note une corrélation négative entre l'âge et la cadence du vibrato et une corrélation positive entre l'étendue et l'âge. Plus le chanteur vieillit (au-delà de 40 ans), plus la cadence est lente et l'étendue devient large.

Dans leur étude, Wapnick et al (1997)^[340] tentent d'identifier les critères sur lesquels les juges se basent pour évaluer une production vocale en soliste. Douze critères sont généralement acceptés :

1. un vibrato approprié,
2. une résonance (ring of the voice),
3. la couleur/la chaleur,
4. l'intensité,
5. l'étendue dynamique,
6. la régularité du registre,
7. l'efficacité dans la gestion du souffle,
8. la flexibilité,
9. le degré de liberté dans l'étendue vocale,
10. la précision de l'intonation,
11. le legato,
12. la diction.

Lors de l'évaluation de ces critères, le jury d'experts, constitué de professeurs de chant, montre un accord inter-juges modéré et un accord intra-juges cohérent. Les auteurs soulignent les difficultés rencontrées dans l'évaluation sur le plan de la fiabilité et estiment le jugement vocal plus délicat que le jugement instrumental. C'est l'importance du timbre qui rend l'évaluation vocale plus difficile. Les professionnels du chant ne trouvent pas de consensus dans ces situations d'évaluation. Le manque de critères objectifs ne signifie pas qu'il est impossible d'aboutir à un consensus général, mais souligne que l'évaluation perd de sa signification.

Van den Berg et Vennard (1959)^[332] ont proposé 22 critères pour une production vocale "souhaitée" et 23 critères pour une production vocale défectueuse [tableau 19].

Tableau 19 : Critères pour une production vocale souhaitée et une production vocale défectueuse selon Van den Berg et al^[332].

Production vocale souhaitée	Production vocale défectueuse
Liberté (<i>freedom</i>)	Lisse (<i>smothered</i>)
Absence d'interférence (<i>lack of interference</i>)	Serrage (<i>straining</i>)
Anneau (<i>ring</i>)	Voix étranglée (<i>throttled</i>)
Formant du chanteur (<i>singer's formant</i>)	Inefficacité (<i>inefficient</i>)
Précision de l'intonation (<i>intonation accuracy</i>)	Tension (<i>tense</i>)
Résonance (<i>resonance</i>)	Fuite (<i>leaky</i>)
Timbre (<i>timbre</i>)	Souffle (<i>breathy</i>)
Couleur (<i>color</i>)	Hurllement (<i>yelling</i>)
Brillance (<i>brilliance</i>)	Forçage (<i>forced</i>)
Puissance (<i>power</i>)	(<i>reaching</i>)
Intensité (<i>intensity</i>)	Peu profonde (<i>shallow</i>)
Focus (<i>focus</i>)	Blanche (<i>white</i>)
Corps (<i>body</i>)	Voix engorgée (<i>throaty</i>)
Profondeur (<i>depth</i>)	Voix fortement serrée (<i>clutched</i>)
Localisation du placement vocal (<i>sensation of height or high placement</i>)	Voix trop couverte (<i>too covered</i>)
Velours (<i>velvet</i>)	Voix pincée (<i>pinched</i>)
Qualité de flottement (<i>floating quality</i>)	Son vibrant (<i>twangy</i>)
Moelleux (<i>mellowness</i>)	Voix nasale (<i>nasal</i>)
Clarté et pureté de la production des voyelles (<i>clarity and purity of vowel production</i>)	Klaxon (<i>honky</i>)
Vibrato approprié (<i>appropriateness of vibrato</i>)	Mugissement (<i>hooty</i>)
Coordination pneumophonique (<i>efficient use of breath support</i>)	Voix étalée (<i>spread</i>)
Flexibilité (<i>flexibility</i>)	Voix sourde et diffuse (<i>dull and diffused</i>)
	Manque d'intensité (<i>lack of intensity or focus</i>)

Ils rapportent que les experts ne sont pas nécessairement d'accord sur la signification technique des paramètres, mais qu'ils trouvent un consensus sur l'aspect musical. Même les experts éprouvent des difficultés dans l'évaluation vocale. Leurs résultats montrent que ces items sont corrélés avec l'évaluation globale de la voix (G) et sont donc importants dans la détermination de l'impression globale. Les items ne sont pas liés à l'âge ($r = -0.18$), ni à l'expérience dans l'enseignement ($r = -0.21$), et pas davantage à l'expérience de l'évaluation ($r = -0.18$). Les auteurs concluent que l'habileté à évaluer de façon cohérente n'est pas enseignée dans les cours de chant.

Les 12 items sont significativement corrélés entre eux. De l'analyse factorielle ressorte 3 facteurs : la qualité intrinsèque (couleur/chaueur, vibrato approprié, étendue dynamique, intensité), la qualité d'exécution (flexibilité, régularité dans le registre, le degré de liberté sur toute l'étendue, l'efficacité pneumophonique, la précision, le legato) et la qualité articulatoire (la diction). Ces trois facteurs rendent compte de 76.2% de la variance.

Sundberg et al (1993)^[311] soulignent plusieurs aspects perceptuels importants. Les qualités perceptuelles de base dépendent du timbre incluant la voyelle et la qualité vocale, de la hauteur tonale et de la puissance. Le timbre est fermement lié à la voix et à la fréquence de résonance du tractus vocal ou formant.

Eklom et al (1998)^[71] proposent à 7 juges experts d'évaluer 5 critères, la résonance, la clarté, la couleur, le vibrato, un grade global, sur une échelle graduée de 0 à 7. Pour chaque paramètre, les auteurs rapportent une fiabilité inter-juges élevée (>0.80). Ils observent que le vibrato influence l'évaluation de l'ensemble de la production. Ils remarquent que si les juges ne perçoivent pas de vibrato dans la demi-seconde, ils l'évaluent avec un score plus bas et que l'absence de vibrato a un effet préjudiciable sur l'évaluation de la couleur. L'évaluation vocale d'un segment isolé de son contexte musical peut faciliter la justesse du jugement dans la manière de se fixer sur certains critères de la qualité vocale en évitant les biais tels que l'exécution technique ou l'interprétation musicale. Les paramètres acoustiques qui affectent l'évaluation de la qualité vocale ne sont en général pas indépendants l'un de l'autre. Il est difficile de les isoler perceptuellement.

Nous proposons dans notre étude une échelle composée de 6 paramètres pour l'analyse du vibrato : non gênant, harmonieux, agréable, régulier, stable et timbré ; nous en avons éprouvé la fiabilité. Dix juges experts ont évalué le vibrato de 54 échantillons vocaux.

3.2.3.2. Matériel et Méthode.

- Les juges

Nous avons choisi 10 juges experts, huit chanteurs d'opéra et deux pianistes accompagnatrices du cours de chant d'opéra. Ils ont évalué les échantillons vocaux de 9 chanteurs. Nous pouvons les considérer comme experts dans la mesure où ils professent au Conservatoire et mènent depuis plus de dix ans une carrière de soliste sur les scènes européennes. Ils participent depuis de nombreuses années à l'évaluation et à la formation de futurs chanteurs professionnels.

Le mois précédant l'évaluation perceptuelle, les juges ont reçu une cassette audio reprenant des exemples de vibrato, trémolo, et chevrottement dans toutes les tessitures vocales, produits par des chanteurs ayant mené une carrière de 1920 à 1950. Pour une question de déontologie, nous trouvions inadéquat d'établir une classification de chanteurs contemporains. La cassette a été constituée par un chanteur dont l'oreille a déjà été éprouvée dans une étude antérieure^[207]. Les juges pouvaient nous contacter pour discuter des échantillons s'ils éprouvaient des difficultés à comprendre le classement effectué, ou s'ils désiraient de plus amples informations sur les catégories (vibrato, trémolo, chevrottement).

- Les chanteurs

Les 9 chanteurs, dont 5 femmes et 4 hommes, sont issus de différentes académies et conservatoires. Leur degré de compétence s'étend de faible à bon et ne sont pas toujours fonction de leurs niveaux d'étude. Leur tessiture est déterminée dans les limites du possible (tableau 20).

Ils produisent trois échantillons différents : un [a] non vibré, un [a] vibré, et un morceau imposé de 18 mesures [figure 18] adapté dans trois tessitures, soprano, alto/baryton, basse.



Figure 18 : Partition du morceau imposé pour soprano composé par G. Chapelier (non publié).

Les enregistrements ont été effectués à l'aide d'un enregistreur numérique D.A.T. (Digital Audio Tape) T.D.C.-D100, Sony, Japan et d'un micro SHURE SM58L-C, Mexico.

Tableau 20 : Degrés de compétence, tessiture et niveaux d'étude des 9 sujets chanteurs.

Chanteurs	Sexe	Degrés compétence ¹⁷	Tessiture	Niveaux d'étude
1	F	Faible	Indéterminée	Étudiante débutante
2	F	Faible	Mezzo-soprano	Étudiante débutante
3	H	Faible	Ténor	Étudiant confirmé
4	H	Moyen	Basse Baryton	Étudiant confirmé
5	H	Moyen	Ténor	Étudiant confirmé
6	H	Moyen	Ténor	Professionnel
7	F	Bon	Soprano	Semi-pro
8	F	Bon	Soprano	Professionnelle
9	F	Bon	Soprano	Professionnelle

¹⁷ Le **degré de compétence** des chanteurs résulte des résultats d'une première évaluation perceptuelle effectuée dans une étude antérieure^[207].

- Tâches à évaluer.

1. Le choix catégoriel.

Les juges devaient dans un premier temps établir un jugement catégoriel de l'échantillon perçu. Quatre catégories sont proposées : le chevrottement, le vibrato, le son blanc, et le trémolo. Elles sont définies comme suit :

Le **chevrottement** est un changement de hauteur tonale avec périodicité et amplitude larges et inégales.

Le **vibrato** consiste en une fluctuation périodique de la fréquence fondamentale, de l'intensité et du timbre du son.

Le **son blanc** est un son sans vibrato.

Le **trémolo** ne possède pas la régularité du vibrato. Sa pulsation est trop rapide et irrégulière en amplitude.

La catégorie **autre** permettait aux évaluateurs d'ajouter un autre type d'oscillation ou d'y classer les voix qu'ils ne pouvaient déterminer.

Les juges devaient porter leur attention uniquement sur l'oscillation et devaient faire abstraction de l'interprétation musicale.

2. L'évaluation des paramètres.

Dans un second temps, les juges devaient évaluer six paramètres : non- gênant (**NG**), harmonieux (**H**), agréable (**A**), régulier (**R**), stable (**S**), timbré (**T**), sur une échelle à 5 points (0 à 4). Les paramètres ont été choisis avec l'avis de trois chanteurs professionnels selon la technique du "brain-storming". Chaque chanteur devait établir une liste de paramètres qualifiant le vibrato. Seize termes étaient communs aux trois chanteurs. De cette liste, nous avons éliminé les termes redondants pour n'en conserver que huit [tableau 21]. Cette échelle a été pré-testée sur ces trois chanteurs. Les termes "mature" et "précis" ont été éliminés. "Mature" ne recouvrait pas la même réalité perceptuelle pour les trois chanteurs. "Précis" était assimilé à "régulier".

Tableau 21 : Paramètres choisis en fonction des tris.

1 ^{er} essai	1 ^{er} tri	2 ^{ème} tri
Agréable	Non- gênant	Non- gênant
Chaud	Harmonieux	Harmonieux
Coloré	Agréable	Agréable
Régulier	Mature	Régulier
Plaisant	Précis	Stable
Souple	Régulier	Timbré
Fin	Stable	
Stable	Timbré	
Mature		
Soutenu		
Timbré		
Non- gênant		
Subtil		
Harmonieux		
Précis		
Léger		

3.2.3.3. Résultats.

Pour la concordance intra-juges, nous avons voulu savoir si chaque membre du jury d'écoute (n=10) reproduisait régulièrement le même jugement, tant pour le type d'oscillation à déterminer (vibrato, chevrottement, trémolo, autre) que pour l'évaluation des paramètres non gênant, harmonieux, agréable, régulier, stable et timbré (NG, H, A, R, S, T).

Le tableau 22 indique le pourcentage de concordance et l'index Kappa pour le type d'oscillation. Selon l'échelle de Landis et Koch, un seul juge montre une fiabilité "substantial" (0.674), 7 une fiabilité "moderate" (comprise entre 0.4 et 0.6) et 2 une fiabilité "fair" (comprise entre 0.2 et 0.4).

Tableau 22 : Pourcentage de concordance intra-juges pour le type d'oscillation.

Juges	% concordance	Kappa	SE Kappa
1	59	0.470	0.133
2	82	0.463	0.125
3	67	0.262	0.118
4	59	0.563	0.157
5	63	0.472	0.123
6	59	0.674	0.132
7	41	0.512	0.133
8	78	0.457	0.126
9	67	0.549	0.123
10	56	0.384	0.133

Le tableau 23 affiche le résultat de la fiabilité globale, tous juges confondus (N=10). Elle est "moderate" ($k=0.471$).

Tableau 23 : Pourcentage de concordance inter-juges pour le type d'oscillation, tous juges confondus.

	Percentile 25	Médiane	Percentile 75
% concordance	58	61	70
Kappa	0.438	0.471	0.552

Le tableau 24 détaille les valeurs du coefficient de Spearman pour la concordance intra-juges sur les 6 paramètres de l'échelle. Pour 8 juges, les valeurs sont comprises entre 0.825 et 0.942 pour le paramètre (T) (timbré). Pour 9 juges, les valeurs sont comprises entre 0.625 et 0.776 pour les paramètres (NG), (H), (A) (non gênant, harmonieux, agréable). Pour 7 juges, les valeurs sont comprises entre 0.446 et 0.598 pour les paramètres (R) et (S) (régulier et stable). Pour 3 juges, les valeurs sont comprises entre 0.067 et 0.367 pour (R) et (S).

Tableau 24 : Concordances intra-juges pour les 6 paramètres (NG, H, A, R, S, T).

Juges	(NG)	(H)	(A)	(R)	(S)	(T)
1	0.638	0.544	0.635	0.67	0.137	0.719
2	0.517	0.342	0.451	0.131	0.367	0.844
3	0.694	0.844	0.830	0.349	0.267	0.863
4	0.776	0.746	0.713	0.446	0.468	0.879
5	0.653	0.673	0.598	0.613	0.516	0.825
6	0.625	0.681	0.681	0.511	0.508	0.677
7	0.352	0.769	0.630	0.563	0.501	0.915
8	0.535	0.607	0.734	0.316	0.585	0.905
9	0.775	0.860	0.816	0.672	0.466	0.942
10	0.725	0.626	0.709	0.542	0.686	0.896

Le tableau 25 montre le taux médian des 10 juges pour les 6 paramètres. Le paramètre (T) (timbré) affiche un taux de 0.881 significativement supérieur aux paramètres (NG), (H), (A) ($p < 0.01$) dont les taux sont de 0.645, 0.677, 0.695. Ces derniers taux sont également significativement supérieurs aux paramètres (R) et (S) ($p < 0.05$) dont les taux sont de 0.478 et 0.484.

Tableau 25 : Concordances inter-juges (N=10) pour les 6 paramètres.

Paramètres	Percentile 25	Médiane	Percentile 75
NG	0.530	0.645	0.737
H	0.591	0.677	0.787
A	0.622	0.695	0.754
R	0.277	0.478	0.575
S	0.342	0.484	0.533
T	0.798	0.881	0.907

Comme **mesure de variabilité inter-juges**, nous avons utilisé la déviation standard des cotes des différents juges par sujet, tâche et paramètre que nous avons étudiés par l'analyse de variance. Le tableau 26 indique les valeurs de la déviation standard pour les 6 paramètres dans les 3 tâches. Les 4 paramètres (**NG**), (**H**), (**A**), (**R**) affichent des valeurs de déviation standard significativement plus petites pour le morceau imposé (M.I.) que pour le [a] vibré (G : $p = 0.019$, H : $p = 0.021$, A : $p = 0.018$). Le paramètre T montre une variabilité significativement plus petite pour le morceau imposé (M.I.) que pour le [a] vibré ($p = 0.006$), ainsi que pour le [a] non vibré ($p < 0.001$).

La déviation standard du paramètre T est également significativement plus petite pour le [a] vibré que pour le [a] non vibré ($p=0.010$). La variabilité inter-juges du paramètre (**S**) ne diffère pas significativement selon les tâches.

Tableau 26 : Variabilité inter-juges pour les 6 paramètres dans les 3 tâches.

Paramètres	A non vibré	A vibré	Morceau Imposé
(NG)	1.92	2.06	1.51
(H)	1.77	2.07	1.52
(A)	1.70	2.07	1.51
(R)	2.59	2.26	2.08
(S)	2.46	2.11	2.24
(T)	2.84	2.23	1.57

Dans les tableaux 27 A, 27 B sont repris les résultats des déviations standards entre les paramètres tous juges confondus et les résultats toutes tâches confondues. Pour le [a] non vibré, les valeurs des paramètres (**NG**), (**H**), (**A**) sont significativement inférieures à celles des paramètres (**R**), (**S**), (**T**). Pour le morceau imposé (M.I.), les valeurs des paramètres (**NG**), (**H**), (**A**), (**T**) sont significativement inférieures à celles de (**R**) et (**S**). Pour le [a] vibré, il n'y a pas de différences significatives entre les paramètres.

Tableau 27 A : Concordances inter-juges pour les 9 chanteurs dans les 3 tâches tous paramètres confondus (Déviations standards).

Chanteurs	[A] non vibré	[A] vibré	Morceau imposé
1	2.05	1.85	2.15
2	2.12	2.5	2.69
3	2.27	1.65	1.48
4	2.45	1.96	2.04
5	2.41	2.29	1.71
6	1.96	2.56	1.51
7	2.19	2.02	1.62
8	2.60	2.54	1.95
9	2.13	1.83	1.48

Tableau 27 B : Valeurs de p pour les déviations standards significatives entre les 6 paramètres.

Paramètres	[A] non vibré	Paramètres	Imposé
NG vs R	P=0.003	NG vs R	P=0.011
NG vs S	P=0.016	NG vs S	P=0.011
NG vs T	P<0.000	H vs R	P=0.014
H vs R	P=0.004	H vs S	P=0.001
H vs S	P=0.002	A vs R	P=0.011
H vs T	P<0.000	A vs S	P=0.001
A vs R	P<0.000	R vs T	P=0.024
A vs S	P=0.001	S vs T	P=0.003
A vs T	P<0.000		

Nous avons aussi évalué la **concordance des jugements au test et au re-test**. La première évaluation porte sur les jugements attribués au type d'oscillation toutes tâches et tous juges confondus. Les jugements sur le vibrato, le son blanc et le chevrottement sont fidèles au test et au re-test, respectivement à 76%, 70% et 62%. Les jugements sur trémolo et autre le sont à 31% en moyenne (Tableau 28).

Tableau 28 : Types d'oscillation test/re-test. Comparaisons par rang toutes tâches et tous juges confondus.

Catégories	Trémolo	Vibrato	Son blanc	Chevrottement	Autre
Trémolo	42	11	17	14	17
Vibrato	3	76	8	12	1
Son blanc	2	15	70	3	10
Chevrottement	4	12	15	62	8
Autre	8	29	42	0	21

Nous avons ensuite évalué la concordance des mêmes variables pour chaque tâche considérée séparément [tableau 29].

Pour le [a] vibré, 70% des jugements sur le vibrato, le son blanc et 69% pour le chevrottement sont identiques à la première et à la deuxième passation du test. Les jugements portés sur les catégories trémolo et autre le sont successivement à 25% et 30%.

Tableau 29 : Résultats pour la tâche [a] vibré pour l'oscillation tous juges confondus.

Catégories	Trémolo	Vibrato	Son blanc	Chevrotement	Autre
Trémolo	25	0	42	25	8
Vibrato	3	70	7	21	0
Son blanc	4	9	70	13	4
Chevrotement	0	13	13	69	6
Autre	10	10	50	0	30

Pour le [a] non vibré, la meilleure concordance est attribuée au vibrato et au son blanc (successivement C=72% et C=69%), la moins bonne au chevrotement (C=25%). La fréquence de jugement des catégories trémolo et autre est nulle. Par conséquent, la concordance l'est aussi [Tableau 30].

Tableau 30 : Résultats pour la tâche [a] non-vibré, pour l'oscillation tous juges confondus.

Catégories	Trémolo	Vibrato	Son blanc	Chevrotement	Autre
Trémolo	0	25	25	0	50
Vibrato	0	72	16	12	0
Son blanc	2	18	69	0	10
Chevrotement	0	25	25	25	25
Autre	0	50	50	0	0

Pour l'imposé, quatre types de jugements sont concordants aux deux passations (moyenne de C=70%) (trémolo, vibrato, son blanc et chevrotement). Le jugement autre n'est concordant qu'à 33% [Tableau 31].

Tableau 31 : Résultats pour la tâche du morceau imposé pour l'oscillation tous juges confondus.

Catégories	Trémolo	Vibrato	Son blanc	Chevrotement	Autre
Trémolo	60	15	0	10	15
Vibrato	5	85	3	5	3
Son blanc	0	16	68	0	16
Chevrotement	17	0	17	67	0
Autre	17	33	17	0	33

Enfin, nous avons observé la **concordance d'évaluation au test et au re-test de chaque paramètre** pris séparément toutes tâches et tous jugements confondus.

Pour le paramètre timbré, les jugements extrêmes de 0 et 5 sont concordants à 93% et les jugements de 1 et 4 le sont à 60%. Les résultats des autres paramètres (NG, H, A, R et S) ne montrent, pour ce type d'analyse, aucune particularité.

Au vu de ces derniers résultats, nous avons effectué une analyse en composantes principales pour chacune des trois tâches considérée séparément.

Pour le [a] vibré, deux facteurs sont mis en évidence. Le premier comprend les paramètres non gênant, harmonieux et agréable. Le second comprend les paramètres régulier et stable. La variance expliquée par le premier facteur est de 66%. Celle expliquée par le second facteur est de 21% (tableau 32).

Tableau 32 : Analyse en composante principale, paramètres X tâche ([a] vibré).

Paramètres	Facteur 1	Facteur 2
Gênant Test	0.846	-0.2
Gênant R test	0.749	-0.361
Harmonieux Test	0.871	-0.225
Harmonieux R test	0.855	-0.353
Agréable Test	0.847	-0.238
Agréable R test	0.819	-0.383
Régulier Test	0.551	0.701
Régulier R Test	0.658	0.512
Stable Test	0.453	0.730
Stable R Test	0.641	0.506
Timbré Test	0.711	0.062
Timbré R test	0.749	0.041

Les deux mêmes facteurs expliquent 98% de variance pour la tâche de l'imposé. Pour ces deux analyses, seul le paramètre timbré reste isolé des deux facteurs [tableau 33].

Tableau 33 : Analyse en composante principale, paramètres X tâche (morceau imposé).

Paramètres	Facteur 1	Facteur 2
Génant Test	0.902	0.014
Génant R test	0.883	-0.204
Harmonieux Test	0.942	-0.072
Harmonieux R test	0.910	-0.245
Agréable Test	0.929	-0.163
Agréable R test	0.871	-0.325
Régulier Test	0.745	0.454
Régulier R Test	0.667	0.581
Stable Test	0.674	0.612
Stable R Test	0.709	0.546
Timbré Test	0.785	-0.406
Timbré R test	0.703	-0.440

Enfin, pour le [a] non vibré, l'analyse met en évidence trois facteurs. Le premier (NG, H, A) explique 63% de la variance, le deuxième (R et S) 16% de la variance et le troisième (T) en explique 15%. Ces trois facteurs expliquent ensemble 98% de la variance [tableau 34].

Tableau 34 : Analyse en composante principale, paramètres X tâche ([a] non vibré).

Paramètres	Facteur 1	Facteur 2	Facteur 3
Génant Test	0.785	-0.150	0.413
Génant R test	0.845	-0.010	-0.365
Harmonieux Test	0.800	-0.047	0.365
Harmonieux R test	0.808	-0.210	-0.387
Agréable Test	0.791	-0.056	0.385
Agréable R test	0.814	-0.213	-0.339
Régulier Test	0.638	0.593	0.333
Régulier R Test	0.579	0.537	-0.491
Stable Test	0.578	0.631	0.381
Stable R Test	0.627	0.485	-0.408
Timbré Test	0.673	-0.527	0.111
Timbré R test	0.703	-0.537	0.012

3.2.3.4. Discussion.

La concordance intra-juges est moyenne. Les juges sont moyennement fiables et ce pour les deux tâches (type d'oscillation et évaluation des paramètres). La concordance inter-juges est mauvaise. Chaque juge a sa propre définition tant sur le plan catégoriel que sur le plan des paramètres perceptuels.

Les paragraphes suivants reprennent chaque variable et étayent la discussion.

- La qualité de l'enregistrement

Les juges ont évalué 54 échantillons vocaux chantés sur base d'un enregistrement réalisé à l'aide d'un DAT (digital audio tape). Wapnick et al (1997)^[341] rappellent que les enregistrements biaisent la qualité vocale. Les petites voix sont relativement amplifiées et les grandes voix relativement affaiblies. Il aurait été préférable pour les juges d'entendre les chanteurs en direct. Nous désirions éviter le biais visuel qui influence le jugement^[340]. Le direct, nous aurait obligé à bénéficier d'une infra-structure particulière. Aussi, l'enregistrement autorisait le juge à ré-entendre, si nécessaire, les échantillons.

- Les chanteurs

Pour l'enregistrement des échantillons vocaux, nous avons rencontré les chanteurs à leur domicile. Nous adapter ainsi à leurs desiderata permettait d'établir de bons contacts. Leur crainte principale était d'être jugé. Les assurer qu'ils participaient anonymement à l'expérience était une condition sine qua non. Le fait d'être chez eux et de savoir qu'ils pourraient effectuer plusieurs enregistrements de leur production vocale diminuait sensiblement leur stress. Le temps qu'ils ont consacré à l'étude du morceau imposé n'a pu être contrôlé. Au vu des difficultés qu'il présentait, un entraînement aurait certainement été bénéfique pour les débutants.

- Les juges.

Après avoir écouté la cassette d'entraînement, certains juges ont refusé de participer à l'étude. Ils ne se sentaient pas aptes à évaluer le vibrato des échantillons vocaux présentés. Ils éprouvaient des difficultés à extraire le vibrato de la voix globale. Pour eux, cette globalité forme un ensemble comprenant la technique, la justesse, l'expression, l'émotion, ... et c'est de l'adéquation de ces critères que résulte une "belle" voix. Sundberg^[311] explique dans son article sur l'aspect perceptuel du chant que l'oreille n'est pas construite pour disséquer l'échantillon dans un premier temps, mais au contraire pour globaliser les caractéristiques de l'émission. Il s'avère donc très complexe d'évaluer des paramètres qualifiant divers aspects vocaux. L'auteur note que les variations des caractéristiques de la source vocale sont plus importantes dans la parole que dans le chant. Tous les types de phonation ne sont pas exploités dans le chant d'opéra ; c'est le cas de la raucité, des voix qui "craquent", des voix soufflées. Ces caractéristiques sont issues de l'asymétrie cordale, ainsi que d'autres imperfections ou anomalies typiques produisant des étendues variées parmi les non-chanteurs. Le chant d'opéra ne révèle pas ces particularités. Les juges ont d'autant plus de difficultés à quantifier des caractéristiques plus fines telles que la régularité, la stabilité sur une zone de l'échantillon donné. Dans l'étude de Wapnick et al (1997)^[341], un juge lui a renvoyé une évaluation incomplète. Il expliquait qu'il lui était impossible de séparer la production vocale de l'interprétation. Pour ce juge, les aspects musicaux et émotionnels faisaient partie de la technique tant dans les vocalises que dans les arpèges.

Wapnick et al relatent que les juges éprouvaient des difficultés à utiliser un même standard pour tous les chanteurs. Ils estimaient qu'un étudiant ne pouvait pas être évalué de la même manière qu'un professionnel. Aussi, un sujet peu entraîné pouvait se voir gratifier d'un grade 3 ou 4 parce que le juge lui prévoyait un avenir vocal intéressant. Tout comme un sujet bien entraîné pouvait recevoir le même grade parce que le juge estimait qu'il aurait pu faire mieux avec les capacités vocales perçues. A travers son évaluation, le juge parle de sa technique, de ses exigences vocales, de ses attentes (Miller (1986)^[200]). Il pouvait dès lors expliquer les difficultés vocales perçues selon son propre schéma de fonctionnement.

Un juge sur les dix obtient néanmoins une concordance "substantial" pour le type d'oscillation. Son bagage professionnel et son expérience semblent peut être le

distinguer du groupe. Ce juge est pianiste répétitrice à l'opéra et a rencontré de grandes personnalités de ce milieu. Son travail consiste non seulement à faire répéter les professionnels, mais également à porter son attention sur la justesse, le rythme, l'interprétation. La contribution du répétiteur est souvent très appréciée par les artistes. Leurs oreilles s'avèrent particulièrement sensibles aux défaillances vocales, si petites soient-elles. Leur grande connaissance du répertoire en font des connaisseurs avertis.

- Le temps d'écoute de l'échantillon

Quant au temps d'écoute de l'échantillon, le nôtre était de plus ou moins 20' pour chaque émission des voyelles [a] et d'une minute pour le morceau imposé. Trois juges sur les dix auraient aimé une tâche plus longue tout comme deux juges de l'étude de Wapnick et al^[340].

Deux juges de son étude ont aussi critiqué la courte durée (1 min 15) des échantillons à évaluer. Ils argumentaient qu'un échantillon plus long serait nécessaire pour se forger une image claire de la force ou de la faiblesse des chanteurs. Ils se disaient dès lors incapables d'évaluer le score global. Augmenter la durée de l'échantillon impliquerait d'alourdir la tâche et d'arriver plus rapidement à un seuil de saturation.

- Le type de tâche.

Les échantillons vocaux à analyser ont fait l'objet de diverses études, la voyelle est souvent proposée pour l'analyse vocale objective, la lecture pour l'évaluation perceptuelle. Nous avons proposé le [a] vibré, le [a] non vibré, et le morceau imposé afin de voir quelle tâche faciliterait l'évaluation du vibrato pour le jury. Au vu des résultats, le choix de la tâche [a] non vibré ne semble guère judicieux. Nous pensions que cette tâche nous permettrait de distinguer les voix sans vibrato. Les sons émis sans vibrato sont qualifiés de son blanc ou encore de son droit ("straight tone"). Miller 1986^[200] juge ce type de son sévèrement. Il est considéré comme signe d'une immaturité vocale. Le son droit relève d'une incoordination pneumophonique et est induit par des tensions laryngées. Un manque d'accolement cordal provoque également une absence de vibrato. Cette production est apparue complexe pour les

chanteurs. Certains n'arrivaient pas à chanter sans vibrato. Dejonckere^[in 63] dans une étude sur 23 chanteurs d'opéra ayant pour objectif l'examen du vibrato et du son blanc doit éliminer deux sujets de l'étude qui sont incapables de produire un son blanc. Il observe également que les 21 autres chanteurs ne peuvent supprimer complètement le vibrato sur des sons soutenus, et pour respecter la tâche, retiennent leur émission. Il en résultait un son sans vibrato pendant quelques secondes, puis en définitive apparaissait l'oscillation. Ekholm et al 1998^[71] stipulent qu'un juge qui ne perçoit pas un vibrato dans la demi-seconde évalue défavorablement la couleur de la voix. L'absence de vibrato aurait donc un effet préjudiciable sur l'évaluation de la couleur. Dans ce cas, il se peut que des chanteurs bien notés sur le [a] vibré et le morceau imposé voient leur score fortement diminué sur le [a] non vibré. Par contre, le morceau imposé s'avère plus efficace. En cela, nous ne rejoignons pas Ekholm et al qui observent qu'un segment isolé de son contexte musical peut aider l'évaluateur. Les auteurs précisent que l'évaluateur aurait le temps de se fixer des critères d'analyse sans toutefois être influencé par l'exécution technique ou encore l'interprétation musicale. Nous n'avons pas observé ce phénomène dans notre étude.

Les échantillons vocaux ont été mélangés et tirés au sort pour le test et le re-test.

L'ordre d'apparition est donc aléatoire. Pour l'évaluation d'échantillons de parole, Kreiman et al (1994)^[172] rapportent que les juges stockeraient en mémoire à court terme des exemples typiques pour des qualités perceptives variées. Ils ajoutent que pour développer une série de repères stables dans le cadre de l'évaluation vocale, les auditeurs experts ont besoin de plusieurs années de pratique et d'entraînement. A l'opposé, les auditeurs naïfs manqueraient de modèles internes spécifiques. Ils feraient référence à leur propre modèle de voix. Ceci plaide en faveur d'un jury d'experts. Pour l'évaluation d'échantillons chantés, nous avons choisi un jury d'experts. Une concordance intra-juges moyenne et inter-juges mauvaise, malgré le choix d'un jury expert, corrobore certaines études. Les juges ont une réalité perceptuelle propre de chacun des paramètres et ne se rencontrent pas en un consensus unanime. L'entraînement avec cassette ne semble pas avoir contribué à harmoniser et fixer des points de repères précis concernant le vibrato.

Certains juges se sont dits influencés par l'ordre des productions, bien que nous ayons pris la précaution de leur faire entendre les 5 premiers échantillons avant de commencer l'évaluation. Leur avis rejoint l'étude de Wapnick et al.

Si nous observons le pourcentage de concordance aux types d'oscillation pour le juge 6, c'est surtout pour les catégories "vibrato" (76%), "son blanc" (70%) et "chevrotement" (62%) que le test/re-test est fidèle. Les catégories "autre" et "trémolo" sont moins utilisées particulièrement pour les tâches [a] vibré et non vibré. "Autre" a été choisi par les juges pour les émissions instables, incluant plusieurs types d'oscillations ou des qualités de timbre nasillant voire engorgé. Le "trémolo" a, par contre, retrouvé un peu d'intérêt pour les jugements portés dans le morceau imposé. Les juges éprouvaient des difficultés à distinguer trémolo et chevrotement. Certains même n'étaient pas d'accord avec les définitions suggérées et au cours de l'évaluation demandaient des explications supplémentaires. Le terme de chevrotement semblait regrouper tant les cadences de fréquences et d'amplitudes lentes que rapides. Il faut avouer que la terminologie précise était, jusqu'il y a peu, l'apanage des seules revues scientifiques. Elle commence à s'imposer progressivement dans les cours de chant au prix de multiples discussions. La confusion reste encore très présente. Pour les tâches [a] vibré et [a] non vibré, les catégories "vibrato" et "son blanc" sont plus fréquemment employées. Le morceau imposé quant à lui, comme la tâche de lecture dans l'évaluation perceptuelle vocale, semble plus adéquat.

L'analyse factorielle révèle trois facteurs. Un premier groupe comprend les paramètres de "non-gênant/gênant", "harmonieux/dysharmonieux", "agréable/désagréable", groupe qui pourrait constituer un **facteur esthétique** de la voix ; le deuxième groupe "régulier/non régulier", "stable/instable" pourrait constituer un facteur **physico-acoustique**, et le troisième le timbre, un facteur global de la qualité du vibrato. Le paramètre "timbré/détimbré" récolte un taux élevé de concordance pour 8 juges sur 10. Son **aspect global** le distingue des autres paramètres. On pourrait l'assimiler à la graduation de l'échelle G.R.B.A.S.. De plus, le timbre fait partie du vocabulaire quotidien tant de l'enseignant que de l'élève. Ce terme recouvre non seulement une réalité perceptuelle, mais également pallesthésique et kinesthésique. Dès lors, il devient un paramètre concret qui correspond à des sensations physiques internes précises.

3.2.3.5. Conclusion.

Une concordance intra-juges moyenne et inter-juges mauvaise, nous permet néanmoins de conclure que le morceau imposé est la tâche qui récolte le plus de concordance et ce, particulièrement pour le paramètre timbré. Les catégories "vibrato", "son blanc" recouvrent une réalité perceptuelle plus précise que les catégories "trémolo" et "chevrotement" qui méritent d'être plus largement explicitées. Une comparaison des résultats aux mesures de déviations standards révèlent trois groupes de paramètres. L'analyse factorielle corrobore ces résultats en révélant trois facteurs principaux. Les paramètres "non-gênant/gênant", "agréable/désagréable", "harmonieux/dysharmonieux" peuvent former un facteur esthétique ; les paramètres de "régulier/irrégulier" et "stable/instable" un facteur physico-acoustique ; et le paramètre "timbré/détimbré" un facteur d'évaluation globale.

chapitre 4

Les mesures objectives

Ce chapitre est réparti en 3 parties. La première concerne les mesures objectives réalisées "manuellement", avant que les laboratoires ne disposent de programmes vocaux adéquats. Cette partie détaille le temps maximum phonatoire mesuré à l'aide d'un chronomètre, la capacité vitale calculée à l'aide d'un spiromètre, le quotient phonatoire résultant des mesures précédentes et le phonétogramme réalisé à l'aide d'un clavier et d'un décibel-mètre. A l'heure actuelle ces mesures sont intégrées aux logiciels et facilitent le calcul du rendement vocal. La deuxième partie reprend un nombre de programmes informatiques disponibles sur le marché européen et en décrit succinctement les possibilités. La troisième détaille les logiciels employés dans notre recherche. Nous discutons de l'analyse spectrale et de son application dans le domaine de la phoniatrie. Nous définissons les mesures réalisées par le logiciel M.D.V.P. (Multi-Dimensional Voice Program) soit :

- les mesures de la fréquence fondamentale,
- les mesures de perturbation fréquentielle à court terme,
- les mesures de la perturbation de l'amplitude à court et à long terme,
- les mesures de ruptures vocales,
- les mesures des composantes sub harmoniques,
- les mesures liées à l'irrégularité vocale,
- les mesures liées au bruit,
- les mesures de tremblements.

Les mesures calculées par les stations E.V.A. 1 et 2. y sont également décrites soit :

- les mesures de la fréquence,
- les mesures de l'intensité,
- les mesures du rapport signal/bruit,
- la mesure aérodynamique de fuite glottique,
- la mesure du temps maximum phonatoire,
- la mesure de l'efficacité glottique ou pression sous-glottique estimée,
- la spirométrie,
- la nasalité,
- le phonétogramme

4.1. LES MESURES OBJECTIVES MANUELLES.

4.1.1. Descriptions des paramètres

L'évaluation vocale aérodynamique est définie comme la mesure de pression d'air et de flux aérien en-dessous et au-dessus de la glotte.

La détermination du Temps Maximum Phonatoire (TMP), de la Capacité Vitale (CV) et du Quotient Phonatoire (QP) est une partie de l'évaluation aérodynamique. Le TMP est une mesure d'habileté phonatoire, et la CV une mesure de la fonction pulmonaire. Les deux sont nécessaires pour la mesure du QP. Ces 3 paramètres sont des mesures non invasives nécessitant un chronomètre et un spiromètre. Les mesures de fuite glottique, de débit moyen produit (mean flow rate : MFR) et de débit d'air buccal peuvent être obtenues sur divers programmes logistiques.

4.1.1.1. Le temps maximum phonatoire

Le TMP est le temps le plus long d'une voyelle en phonation soutenue (le plus souvent [a], quelquefois [i] ou [ou]) immédiatement après avoir pris une inspiration maximale.

Raes (1996)^[242] résume les difficultés rencontrées lorsqu'il s'agit de comparer les valeurs rapportées par différents auteurs. Un manque d'uniformité dans la procédure d'exécution de la tâche modifie les résultats. Le type de voyelle utilisée, le nombre d'essais de TMP par sujet, la présentation de la tâche, les instructions données, l'encouragement fourni au sujet, ainsi que la position du sujet, la phonation avec ou sans embouchure influent sur les résultats de ce paramètre (Ptacek et al(1963)^[238], Dobinson et al (1993)^[66], Hirano (1981)^[122]).

Le sexe, l'âge déterminent largement les valeurs de TMP. Les hommes soutiennent le son plus longtemps que les femmes. Entre 3 et 17 ans le TMP croît et au-delà de 65 ans, il diminue et devient inférieur à celui de jeunes adultes.

Hirano et al (1968)^[121] ont déterminé les valeurs moyennes de TMP pour 1563 sujets répartis en 22 catégories de voix pathologiques organiques et fonctionnelles. Pour deux de ces catégories, la paralysie uni et bilatérale des cordes vocales, la valeur moyenne du TMP était inférieure à 6 secondes (5.9 et 5 secondes respectivement). La valeur moyenne pour tous les autres troubles était supérieure à 13 secondes, s'étendant de 13.5 secondes pour les patients atteints d'un cancer laryngé type 2, à 18.6 secondes pour ceux présentant une laryngite aiguë. En dépit de ces différences observées dans la moyenne du TMP entre la paralysie des cordes vocale et les autres pathologies, Hirano et al ont conclu que ces troubles ne pouvaient être différenciés sur la base seule du TMP à cause du chevauchement des valeurs entre ces différents groupes.

Dans le cas de polypes et de paralysies de cordes vocales, le TMP se révèle négativement corrélé avec la taille de la béance glottique. Cette mesure peut être utile pour contrôler le progrès du traitement suivi.

Tableau 35 : Valeurs de temps maximum phonatoire rapportées par Raes^[242].

Auteurs	Années	Âges	Sujets	Valeurs Hommes	Valeurs Femmes
Ptacek et al ^[238]	1963	17-41	80	24.7 (6.07)*	16.8 (5.13)*
Ptacek et al ^[239]	1966	18-39	58	24.6 (6.7)*	20.9 (5.7)*
Hirano et al ^[121]	1968		50	34.6	25.7
Dobinson et al ^[66]	1993	20-59	76	22.2 (9.2)*	18.4 (7.3)*

Légende : *déviations standards entre-parenthèses.

4.1.1.2. La capacité vitale

C'est le volume d'air mesuré à la bouche entre une inspiration et une expiration maximales. Le spiromètre est l'instrument qui permet cette mesure.

Pour obtenir les mesures, le clinicien recourt à différentes méthodes et plusieurs essais. Il peut demander au sujet de souffler calmement ou vivement dans le spiromètre. Il fait répéter la tâche. Dobinson et al^[66] sont partisans de plusieurs essais. Quanjer (1983)^[240] recommande un minimum de 3 essais et un maximum de 8. Il précise que le sujet doit être assis dans une posture droite, la CV variant selon la position corporelle.

D'après Raes, la littérature rapporte peu de valeurs de référence pour ce paramètre. Les valeurs de quotient phonatoire sont plus souvent mentionnées. La CV n'est jamais qu'une mesure de la fonction pulmonaire et non de la phonation ^[242]. Les valeurs rapportées dans le tableau ci-dessous sont considérées dès lors comme des estimations plutôt que des valeurs de référence sur groupes standardisés. Les valeurs de CV du B.S.G.V.D. pour les groupes de dysphonie sont comparées aux données de Ptacek et al^[239].

Tableau 36 : Valeurs de capacité vitale rapportées par Raes^[242].

Auteurs	Années	Sujets	Âges	Valeurs Hommes	Valeurs Femmes
Ptacek et al ^[239]	1966	62	18-39	4800 cc (600)*	3500 cc (600)*
Zemlin ^[362]	-	-	-	4600	3100
Greene ^[102]	-	-	-	4000	3000

Légende : cc : centilitre ; *déviations standards entre-parenthèses.

Si la quantité d'air est extrêmement limitée, cela peut contribuer aux difficultés vocales de n'importe quels locuteurs et particulièrement des locuteurs professionnels et des chanteurs. Des valeurs anormalement basses de CV sont plus souvent retrouvées dans des populations phoniatriques que des populations normales. La CV n'est pas une fonction diagnostique : elle ne donne pas directement d'indications pour le degré de la pathologie vocale.

4.1.1.3. Le Quotient Phonatoire.

Le quotient phonatoire est une mesure approximative de l'air consommé durant la phonation par unité de temps. La valeur est obtenue en divisant la CV par le TMP. Cette valeur augmente dans les dysphonies vocales.

D'un auteur à l'autre (Dobinson et al^[66] ; Hirano et al^[121]), les valeurs observées diffèrent autant que les valeurs de CV et de TMP, le QP découlant directement de ces deux mesures. Les divergences s'expliquent d'autant mieux que les uns multiplient le nombre de tests, les autres ne réalisant qu'une seule fois l'examen. Les valeurs de QP

augmentent avec les troubles vocaux. Hirano et al (1989)^[125], dans une étude plus récente, constatent aussi des valeurs moyennes plus élevées dans les cas de paralysie de cordes vocales unilatérales et bilatérales. Ils trouvent 387 ml/s et 372 ml/s respectivement, alors que les valeurs moyennes de QP dans d'autres troubles sont comprises entre 151ml/sec dans les laryngites aiguës et 244 ml/sec dans les traumatismes laryngés. Les valeurs de QP, cependant, n'ont pas permis de faire de diagnostic. Dans les cas de polypes et de paralysies de CV , ils observent une corrélation entre le QP et la taille de la béance glottique.

Tableau 37 : Valeurs de quotient phonatoire rapportées par Raes^[242].

Auteurs	Années	Sujets	Valeurs Hommes	Valeurs de Femmes
Hirano ^[121]	1968	50	145	137
Dobinson et al ^[66]	1993	76	269 (88)*	233 (82)*

Légende : *déviations standards entre-parenthèses.

De cette revue, il ressort que le TMP et le QP paraissent avoir des rôles similaires en clinique. Ces mesures sont utiles à la fois pour déterminer le degré de désordre vocal et, pour contrôler leur évolution, mais n'offrent pas de possibilités diagnostiques.

4.1.1.4. Le phonétogramme.

Heylen (1996)^[118] rapporte la définition du phonétogramme que donne Schutte (1980)^[279].

L'étendue fréquentielle sur l'axe horizontal d'un phonétogramme comprend la fréquence la plus haute et la plus basse auxquelles les cordes vocales sont capables de vibrer. L'étendue en intensité reportée sur l'axe vertical comprend l'intensité la plus faible et la plus forte. La hauteur de l'échelle comporte 5 octaves et l'étendue de la pression sonore s'étend de 40 à 120 dB.

Lors de l'examen, le sujet produit la voyelle [a] à 30 cm du micro. Le niveau de bruit de la pièce doit être inférieur à 40dB, et la voix doit être physiologiquement

acceptable. Le phonétogramme sert à mesurer les possibilités de la voix chantée, pour le diagnostic des troubles vocaux et l'évaluation des résultats de la thérapie.

L'enregistrement du phonétogramme peut être classique ou manuel, semi-automatique ou automatique. Les instruments de mesure employés dans la méthode classique ou automatique sont un générateur de son et un décibelmètre. Le thérapeute joue ou, chante un ton à une certaine hauteur et demande au patient de produire le même son aussi doucement et aussi fort que possible. Les fréquences sont mesurées dans un ordre ascendant et subséquemment dans un ordre descendant. Elles sont notées par $\frac{1}{2}$ tons. La mesure de pression sonore et de la fréquence de base sont mesurées et imprimées sous forme d'étendue vocale. L'examen est réalisé en 10 à 20 minutes. Le sujet doit donc être capable d'imiter correctement le son produit, le thérapeute de juger de la justesse de la fréquence. Ces exigences constituent un désavantage de la méthode.

Obtenues par ordinateur, les mesures sont enregistrées automatiquement. Suite à la production d'un [a] à hauteur et intensité demandées, un contour rigide est enregistré. L'affichage à l'écran constitue un feedback visuel qui facilite considérablement l'examen pour le patient non entraîné musicalement et le thérapeute sans habileté musicale. Le contrôle de la fréquence qui prévoit des écarts d' $\frac{1}{2}$ tons, est primordial. Le développement du phonétogramme par ordinateur a encouragé son utilisation pour des sujets atteints de troubles vocaux variés. Le phonétogramme serait une aide à la description tant des capacités que des difficultés vocales. Heylen (1996)^[118] rapporte que les auteurs ont souvent fait une distinction entre la production physiologique et la production musicale que le phonétogramme décrit. Les formes sont plus représentatives des limites physiques de la production vocale. Dès lors l'examineur ne prend en compte que le son vocal "acoustiquement" acceptable.

Tout comme celles de la capacité vitale, le quotient phonatoire et le temps maximum phonatoire, les valeurs obtenues au phonétogramme seraient plus influencées par le manque de standardisation de la technique que par la variabilité des capacités vocales des différents sujets^[43]. Coleman (1993)^[43] décrit une série de variables qui influencent les résultats obtenus au phonétogramme, soit la qualité vocale du sujet, les instruments et l'environnement. Schutte et al (1984)^[280] recommandent les sons [a], [u] et [i] tandis que Gramming et al (1988)^[101] se limitent

au son [a]. Gaufin et al (1988)^[86] n'adoptent les fréquences élevées que sur une phonation forte. Ils notent l'intensité la plus forte sur la fréquence la plus élevée.

Durant les productions sonores graves et médium, l'effet des différentes voyelles peut être minimisé parce que le fondamental contient le plus d'énergie spectrale. La production du niveau de pression sonore dépend principalement de l'amplitude de la Fo pour sa limite la plus basse, et sur les tons les plus hauts du modèle vocal pour sa limite supérieure. Coleman^[43] exige que la production vocale soit maintenue au minimum deux secondes, la durée du son influençant clairement le niveau de pression sonore. La variation du niveau sonore de pression à court terme est de +/- 5 dB durant une phonation légère. Celle du niveau de pression sonore à long terme est de +/- 2.4 dB pour une phonation légère et de 3.4 dB pour une phonation forte. La variation de fréquences est de 2 tons ½ à court terme. Les auteurs s'accordent sur le fait que des variations de 2 à 3 dB sont normales pour des essais successifs et qu'un ton produit est acceptable si la déviation maximale est de +/-3% de la fréquence recherchée. Lors du test, les fréquences hautes ne sont pas répétées pour éviter toute fatigue vocale.

Afin d'obtenir le meilleur résultat, Coleman conseille l'échauffement vocal avant de faire passer le phonétogramme. Il rappelle que le moment de la journée peut influencer la production vocale. Un autre facteur est le degré de contrôle du souffle des non-chanteurs. D'autre part, un certain degré d'ouverture de la bouche est nécessaire pour obtenir des résultats optimum ; or les chanteurs entraînés modifient leur aperture vocale et par la même la fréquence de phonation et augmentent l'intensité. D'autres facteurs peuvent agir : le flux d'air à travers la glotte peut entraîner une perte d'intensité de 15 dB ; un vibrato donne généralement des modulations de fréquence et d'amplitude qui peuvent faire varier la hauteur de 0.5 à 2 tons ½ et l'intensité de 2 à 10 dB. Pabon (1988)^[225] observe qu'il est nécessaire pour obtenir de bons résultats d'encourager le patient à sonoriser de façon maximale.

Les changements soudains dans le profil du phonétogramme reflètent les passages de registres ou les zones des formants de la voyelle. Les auteurs s'accordent pour reconnaître que cette méthode diffère des examens conventionnels et explique une différence majeure d'interprétation entre les examinateurs qui emploient le phonétogramme dans leur pratique clinique, et les autres.

Certains praticiens l'emploient dans le cadre des voix normales et chantées (Schutte et al (1983)^[280], Awan (1991)^[12], Morris et al (1995)^[203]) ou uniquement

chantées (Waar (1968)^[338]). D'autres le recommandent pour déterminer la capacité de fonctionnement tant pour les voix normales que pathologiques (Klingholtz et al (1989)^[165], Hacki (1993)^[106]). Coleman (1993)^[43] ajoute qu'il aide à vérifier la maturité des possibilités vocales et à faire la distinction entre les voix entraînées et non entraînées. Il précise que la méthode peut aussi donner des informations valables en cas de dysphonies dysfonctionnelles. Schultz (1975)^[278] examine les troubles fonctionnels et organiques, compare les résultats intra-individuels, assure le suivi vocal et les capacités vocales. Il réalise des comparaisons inter-individuelles entre la voix du sujet et les valeurs de référence. Pour les voix normales, il étudie l'impact de l'entraînement sur le phonétogramme chez les enfants et les adultes. Pedersen (1993)^[228] réalise des profils vocaux dans un chœur de filles et de garçons en période de mue. Heylen (1996)^[118] vérifie les capacités vocales des enseignants masculins à l'école primaire. Il conclut que l'influence du stress, les modifications post-thérapeutiques, la détermination des registres, l'étude des paramètres en pré et post-chirurgie sont autant de facteurs cliniquement utiles à analyser à l'aide du phonétogramme. Le phonétogramme est complémentaire au diagnostic vocal. Il est plus une réflexion sur la capacité vocale qu'une mesure de la fonction laryngée, mais en combinaison avec les données laryngoscopiques, il peut nous renseigner sur cette aptitude vocale des sujets.

Les caractéristiques du phonétogramme peuvent être schématiquement groupées autour des paramètres suivants : fréquence, intensité, forme et surface, qualité.

L'étendue fréquentielle est exprimée en Hz en ½ ton ou en % par rapport à l'étendue totale. L'examineur peut évaluer le volume vocal complet, la tessiture vocale ou le volume physiologique du registre de tête. Le phonétogramme est aussi une méthode pour déterminer les registres et les transitions de registre, les classer par tessiture vocale, mesurer l'intensité minimale ou maximale, confortable et habituelle.

L'inclinaison est une des formes de paramètre étudié par Hereto^[in 118]. Cependant aucune définition claire, aucune méthode de calcul n'apparaît pour la mesurer. En conséquence, il est difficile de comparer les données d'études variées. Plus encore, l'inclinaison est mesurée en différentes unités :

Gramming et al (1988)^[101] l'expriment en $\frac{1}{2}$ tons par dB, mais dans d'autres études, l'inclinaison est mesurée en dB par $\frac{1}{2}$ tons. Il en est de même pour le calcul de surface.

Pabon (1991)^[226] essaie de mesurer les aspects qualitatifs par les moyennes de la phonétographie. Plus spécifiquement, il examine le jitter, le shimmer, et le facteur crête. Kotby et al (1995)^[162] font une analyse quantitative et qualitative de leurs résultats. Ces analyses qualitatives consistent principalement en des annotations sur la fréquence, l'intensité, et les caractéristiques du phonétogramme.

Heylen (1997)^[119] décompose le phonétogramme en fréquence, intensité et caractéristiques morphologiques. Les 6 caractéristiques fréquentielles sont le pic le plus haut et le plus bas en Hz, l'étendue fréquentielle en Hz, le nombre de $\frac{1}{2}$ tons des registres de tête et de poitrine et l'étendue totale des pics exprimés en $\frac{1}{2}$ tons. Le phonétogramme est ensuite décomposé en trois types d'intensité : l'intensité (dB) douce et forte, l'étendue totale d'intensité exprimée par la différence entre l'intensité forte et l'intensité faible, et l'interaction entre la fréquence et l'intensité reflétée par la morphologie du phonétogramme. Cette morphologie est décrite par les pentes des courbes des frontières supérieures et inférieures illustrée à la figure 19. Hereto^[in 118] propose que l'examineur dessine manuellement une tangente des frontières du registre de tête. Avec les paires X (Hz) et Y (dB) comme indiqué comme à la figure 19, on calcule la pente en dB/1/2 tons.

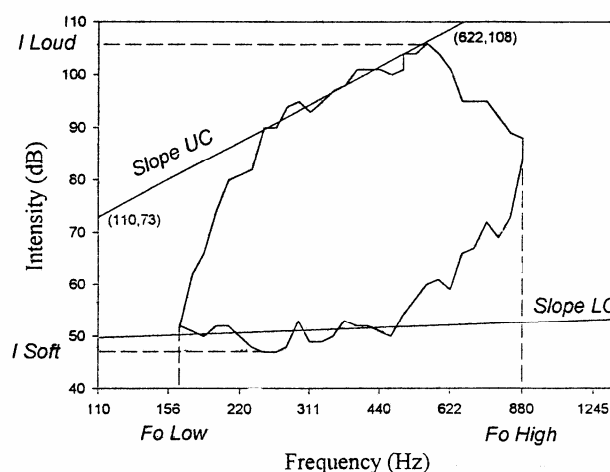


Figure 19.

La revue de la littérature présente plusieurs références de phonétogrammes normalisés. Böhme et al (1995)^[25] rapportent des valeurs de référence pour 112 enfants âgés de 7 à 10 ans. Aucune différence significative en fonction de l'âge n'apparaît dans ce groupe. Les intensités maximales et minimales augmentent avec la fréquence. Les résultats montrent cependant que les garçons chantent plus fort que les filles. L'étendue vocale est de 29 ½ tons pour les deux groupes. A l'origine, la population de Böhme était beaucoup plus large (n=277), du fait que la tranche d'âge était de 5 à 14 ans. Il n'a pu établir de valeurs de référence en dessous de 7 ans, la concentration et la perception nécessaires à la reproduction des fréquences perçues étant impossibles. Les variations inter-sujets chez les plus âgés ne lui ont pas permis de les intégrer au standard.

Tableau 38 : Niveaux de pression sonore en fonction de la hauteur tonale pour des filles (N=67) et des garçons (N=45) âgés de 7 à 10 ans rapportés par Böhme et al (1995)^[25].

Hauteurs tonales (Hz)	Niveaux de Pression mini		Niveaux de pression maxi	
	Filles	Garçons	Filles	Garçons
196	50 (5)	53 (6)	66 (7)	70 (7)
220	50 (5)	53 (4)	72 (5)	73 (7)
247	50 (5)	54 (6)	77 (5)	76 (6)
262	51 (5)	54 (5)	81 (5)	81 (6)
294	54 (5)	54 (5)	82 (5)	81 (6)
330	54 (4)	57 (6)	82 (6)	84 (7)
349	54 (4)	58 (5)	84 (5)	84 (7)
392	57 (4)	59 (6)	83 (6)	85 (7)
440	59 (5)	61 (6)	84 (7)	85 (7)
494	62 (5)	61 (6)	83 (6)	88 (8)
523	62 (5)	65 (7)	83 (7)	86 (8)
587	66 (5)	68 (7)	84 (7)	86 (7)
660	69 (5)	70 (7)	85 (5)	87 (7)
698	70 (6)	73 (6)	86 (5)	88 (7)
784	72 (6)	74 (7)	87 (5)	89 (8)
880	74 (5)	76 (6)	87 (6)	90 (8)
988	75 (6)	78 (8)	87 (7)	91 (8)
1048	74 (6)	78 (8)	88 (6)	89 (8)

Coleman et al (1977)^[41] rapportent des valeurs de référence pour les adultes. Pour les sujets sans pathologie vocale, l'étendue couvre au moins 2 octaves. L'intensité minimale est de 55 dB pour les femmes et 58 dB pour les hommes.

L'intensité maximale est de 113 dB pour les femmes et 117 dB pour les hommes. La dynamique d'amplitude varie en fonction de la fréquence. Elle se situe en moyenne autour de 30 dB. Le passage d'un registre à l'autre se marque par un pincement du champ dynamique particulièrement présent dans les voix non chantées. Dejonckere ^[in 46] relève les anomalies vocales principales pour les dysphonies organiques. Il observe une réduction globale de l'aire, une perte des fréquences aiguës et parfois des fréquences basses. Il rapporte une augmentation des intensités minimales et une diminution des intensités maximales dans les fréquences aiguës. Le phonétogramme peut également être scindé en deux parties. Pour les dysphonies dysfonctionnelles, l'auteur parle de réduction de la dynamique et d'une réduction des deux extrémités de la courbe.

Sulter et al (1995)^[301] réalisent le phonétogramme sur 224 sujets répartis en 4 groupes. 92 femmes et 47 hommes (17-41 ans) constituent le groupe de sujets non entraînés, 43 hommes et 42 femmes (18-75 ans) le groupe de choristes. Ils s'intéressent à la fréquence et à l'intensité. Les résultats sont différents selon le sexe dans les deux populations. Les hommes sont capables de sonoriser à des intensités minimales plus faibles que les femmes. Les choristes masculins montrent une habileté à produire des hauteurs tonales plus basses. Il en résulte un phonétogramme avec des capacités dynamiques plus larges particulièrement sur les intensités minimales. Cela suggère un meilleur contrôle du souffle durant la phonation et de l'habileté à faire vibrer les cordes vocales à une pression sous-glottique basse. A la fin de leur étendue fréquentielle, les sujets non entraînés produisent des intensités plus fortes, tandis qu'une des caractéristiques des choristes est la relative absence de transitions de registre. Les différences entre les deux groupes s'expliquent par des capacités naturelles plus grandes chez les choristes à contrôler leur mécanisme vocal. L'expérience du chanteur est associée à un contrôle neuro-myogénique sur la source vocale.

Neuschaefer et al (1997)^[216] évaluent la production vocale de 40 orateurs professionnels et non-professionnels à l'aide d'un phonétogramme tridimensionnel. Les sujets sont répartis en quatre groupes appariés en âge et en sexe. Sur le phonétogramme en 3 dimensions, l'axe X indique les valeurs de hauteurs tonales, l'axe Y les valeurs d'intensité et l'axe Z, la durée de la phonation. Le concept du

phonétogramme en 3 dimensions a été développé par Frank et al (1983)^[83], Gross et al (1983)^[105], et Lillge (1984)^[183]. Les résultats montrent que les phonétogrammes des professionnels sont plus "volumineux" que ceux des non-professionnels [figure 20].

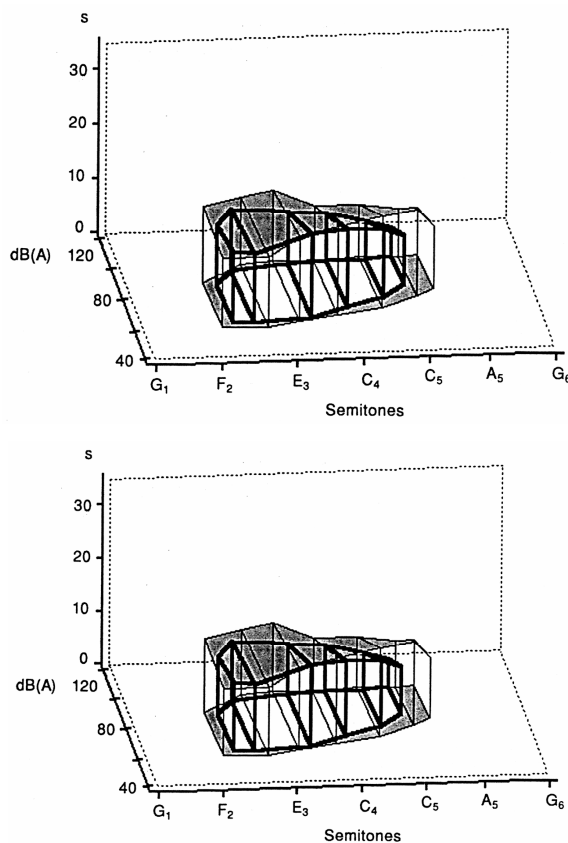


Figure 20 : Phonétogrammes en 3 dimensions d'orateurs professionnels et non professionnels hommes et femmes. Neuschaefer et al (1997)^[216].

Les orateurs non-professionnels montrent une étendue de 29 ½ tons pour les femmes et 32 ½ tons pour les hommes, les professionnels une étendue de 35 ½ tons pour les femmes et 37 ½ tons pour les hommes. Chez femmes les valeurs d'intensité faible diffèrent sur toute la tessiture, celles d'intensité fortes sur les fréquences moyennes et hautes. Les femmes professionnelles obtiennent une durée de phonation plus élevée en intensité faible que les non-professionnelles (11-14 sec. versus 13-19 sec.). Pour les intensités fortes, aucune différence significative n'est rapportée. Chez les hommes, l'effet de l'entraînement vocal est démontré. Les orateurs professionnels

obtiennent des valeurs d'intensité minimale plus faibles et des valeurs d'intensité maximale plus fortes que les non-professionnels et ce sur toute la tessiture ; la durée de la phonation est également supérieure (10-17 sec. versus 13-19 sec.).

Tableau 39 A : Niveaux de pression sonore et temps maximums phonatoires pour des orateurs professionnels et non-professionnels ^[216].

Orateurs	HT(intervalle)	Intensité (dB)		TMP (sec)	
		Faible	Forte	Faible	Forte
Non-professionnels	87-110	58.8	83	14.4	11.9
	123-165	57.1	90.5	3.5	13.3
	175-220	59.3	93.3	16.5	11.1
	250-330	65	94.7	16.3	9.8
	349-440	69.6	95.1	13.6	7.6
	494-523	77.6	97.6	9.6	6.3
Professionnels	82-110	55.4	84	14	15.8
	110-147	54.5	94.8	14.3	14.1
	165-220	57.4	96.1	16.8	14.6
	220-294	60.2	94.4	18	12.3
	329-440	63.2	97.5	19.4	10.9
	440-587	70.1	95.7	17.8	9.7
	659-698	73.8	96.9	13.4	10.4

Tableau 39 B : Niveaux de pression sonore et temps maximum phonatoire pour des oratrices professionnelles et non-professionnelles ^[216].

Oratrices	HT(intervalle)	Intensité (dB)		TMP (sec)	
		Faible	Forte	Faible	Forte
Non-professionnelles	147-196	58.1	77.3	11.1	14.3
	196-262	59.5	88.7	12.1	14.6
	294-392	60.5	89	12.8	12.5
	392-523	67.2	87.4	13.2	10.2
	587-784	74.4	87.6	13.7	9.2
Professionnelles	131-175	51.9	72.9	12.9	16.3
	196-262	52	88.6	15.8	16
	262-349	55	90	18.6	13.2
	392-523	60.8	95.2	18.9	12.3
	523-698	65.9	99.5	19	12.2
	784-1046	72.7	98	14	10.5

L'analyse combinée des 3 paramètres (fréquence, intensité, durée) a montré que les valeurs de production des orateurs professionnels dépassaient celles des non-

professionnels. Les auteurs concluent que cette nouvelle forme de phonétogramme peut être considérée comme un outil utile et pratique dans l'estimation de la capacité vocale.

4.2. L'APPORT DES LOGICIELS.

"Dans le cadre d'un phénomène complexe, multidimensionnel comme la voix, toute mesure isolée est réductrice ; il n'existe aucun paramètre dont la mesure rende compte de toutes les dimensions d'une dysphonie" Giovanni (1998)^[98].

Le développement de logiciels et des notions de travail par ordinateur permet un accès rapide à de multiples données acoustiques et aérodynamiques parmi lesquelles le clinicien a du mal de faire le tri entre ce qui est nécessaire et ce qui est redondant. Ces appareils ont l'avantage de la rapidité d'acquisition de données et de leur stockage. Nous donnons ici un aperçu des logiciels présents sur le marché européen. Nous avons travaillé sur trois modèles. Il convient de sélectionner, en fonction des besoins de chacun, les informations pertinentes et de les intégrer au bilan clinique, à l'auto-évaluation, à l'évaluation perceptuelle afin de déterminer l'attitude thérapeutique et le suivi du patient.

Les firmes **Brüel et Kjaer** (Naerum, Danemark), et **Hewlett Packard** (Palo Alto, Californie) ont créé de nombreux appareils destinés à l'analyse vocale. Leurs champs d'application concernent autant le domaine de l'acoustique que celui de l'électro-acoustique, de la vibration ou encore de la médecine. Leurs analyseurs vocaux fournissent le sonagramme, le spectrogramme. On peut également obtenir les paramètres de fréquence et d'intensité en fonction du temps.

D'autres programmes ont été mis au point par des laboratoires dans le but de répondre aux nécessités de la clinique [annexe 6].

PCLX est un programme comportant deux logiciels PCWAVE et PCPITCH. PCWAVE permet une visualisation de la parole en temps réel et fournit une courbe laryngographique, PCPITCH permet d'obtenir une visualisation du contour de la fréquence fondamentale et une mesure de la période sur la courbe laryngographique.

Phonedit est un éditeur de signal de parole qui permet également sa manipulation et son traitement. Ce programme a été développé au laboratoire de recherche "Parole et langage" de l'université de la Méditerranée. Il permet d'obtenir des mesures d'intensité et de durée, une analyse spectrale et un sonagramme.

Phonos 1 comprend plusieurs logiciels répartis en modules. Deux de ces modules permettent l'édition du signal de parole, une analyse spectrale, l'extraction de la fréquence fondamentale, une visualisation en temps réel de l'intensité, du rythme, de l'articulation des sons fricatifs, le temps maximum phonatoire et une représentation de la durée expiratoire.

Signalysé créé par Keller E, analyse tous les signaux acoustiques (sons, parole, musique,...). Le signal offre une édition du signal de parole, l'extraction de la fréquence fondamentale, une analyse spectrale à long terme, un spectrogramme, un spectre moyenné, le cepstre, et l'étiquetage phonétique.

Spectra plus est un logiciel d'analyse spectrale créé par Cardot G, Peret Ph et Patte J. Il s'applique dans l'enseignement, l'acoustique architecturale, la recherche médicale, l'O.R.L., la logopédie, l'audio-phonologie, la sonorisation professionnelle. On obtient l'enveloppe vocalique (amplitude en fonction du temps), le spectre, le spectre en 3 dimensions, et le sonagramme.

N-UNICE est un logiciel d'analyse acoustique et phonétique de la parole. Il fournit spectre, spectrogramme, édition du signal de parole, courbe mélodique, courbe d'intensité et étiquetage phonétique.

O-WINSNORRI, logiciel d'étude de parole, est destiné à la recherche en phonétique. Il permet l'édition du signal de parole, le spectrogramme, l'analyse spectrale, l'extraction de la fréquence fondamentale, la détection des formants et les étiquetages phonétiques.

C.E.C.I.L. est un logiciel destiné à l'analyse de la voix normale et pathologique. Il permet d'obtenir le signal de parole, le spectre, le spectrogramme, la fréquence fondamentale, la courbe d'intensité, la courbe mélodique, la courbe du jitter et du shimmer.

Dr Speech est un logiciel d'analyse de la parole et d'évaluation de la qualité de la voix. L'analyse quantitative et qualitative de la fonction vocale est réalisée par une édition du signal de parole, une analyse spectrale à long terme, la fréquence fondamentale, l'intensité, le shimmer, le jitter, l'évaluation du signal électroglottographique, le phonétogramme en temps réel et le spectrogramme.

PHONEVAL est un logiciel destiné à l'évaluation des voix pathologiques. Il a été mis au point par Perrin E. Il présente un ensemble d'outils informatisés de traitement du signal de parole par le biais de l'analyse sonographique : l'analyse

spectrale, la détection de la fréquence fondamentale et de l'intensité ainsi que des mesures de perturbations soit le jitter et le shimmer.

SOUNDSCOPE a également un champ d'application très vaste puisqu'il s'applique tant dans le domaine de la clinique que de la recherche , de l'enseignement et de l'industrie. Il fournit l'édition du signal de parole, le spectrogramme, la fréquence fondamentale, le jitter et le shimmer, le spectre, l'enveloppe d'amplitude et le rapport signal/bruit.

4.3. LES APPAREILS UTILISÉS DANS NOTRE RECHERCHE.

Les trois appareils que nous avons utilisés sont l'analyseur dynamique de signal de Hewlett Packard^[117], le "Computer speech lab" de Kay Elemetrics ^[159, 160] et l'Evaluation vocale assistée des Laboratoires Parole et langage de Marseille ^[293, 296]. Nous en faisons une description détaillée aux points suivants.

4.3.1. L'analyseur dynamique de signal de Hewlett Packard (modèle 3561A).

Ce modèle est un analyseur dynamique de signal. Ses capacités d'analyse sont puissantes. Adapté au secteur de la voix, cet instrument est devenu un outil courant d'analyse spectrale. Il offre une étendue fréquentielle d'analyse s'échelonnant de 0 à 100.000 Hz avec une plage dynamique de 80 dB et une précision en amplitude de 0.15 dB.

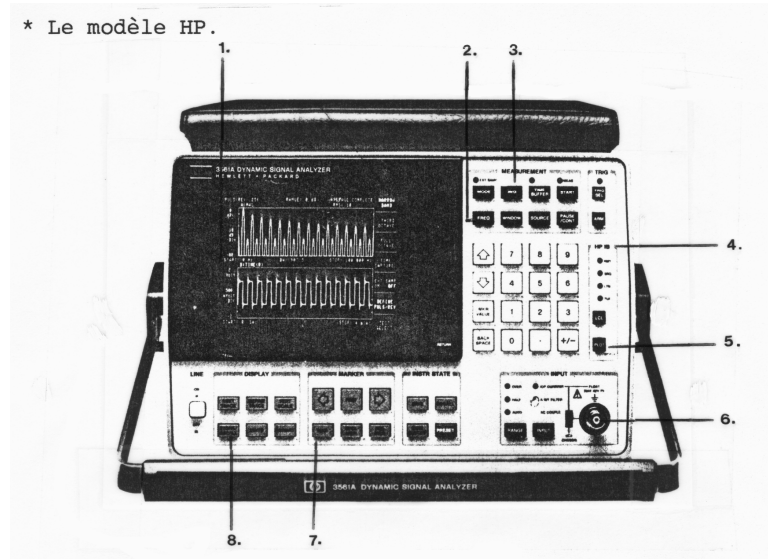


Figure 21 : Analyseur dynamique de signal Hewlett Packard modèle 3561 A¹⁸ ^[117].

¹⁸ Mesures simultanées dans le domaine du temps et de la fréquence.

² Plage de mesure de 100 kHz.

³ "Moyennage" pour améliorer la précision.

⁴ Fonctionnement en programmation HP-IB.

⁵ Tracé numérique directe.

⁶ Dynamique de 80 dB, précision en amplitude de plus +/- 0.15 dB.

⁷ Zone de commande des marqueurs de précision.

⁸ Mémoire permanente : elle stocke 127 traces et configurations.

4.3.1.1. L'analyse spectrale.

La parole résulte des actions du tractus vocal. Toute modification acoustique du signal devrait être le signe d'une modification des organes vocaux.

Le vibrateur produit le son, qui est filtré par le tractus vocal tant en fréquence qu'en amplitude. La muqueuse laryngée absorbe l'onde glottique. Les résonateurs renvoient l'énergie acoustique sous forme de fréquences et d'harmoniques. Pour chaque voyelle, le volume et la forme du tractus vocal changent ; les caractéristiques de résonance sont modifiées, ce qui influence l'amplitude spectrale.

En analyse spectrale, la quantité d'informations d'une émission simple telle que celle d'une voyelle est importante. Beaucoup d'études ont tenté de déterminer quels aspects ont du sens, quels autres peuvent être ignorés dans la recherche de conclusions relatives au comportement productif et perceptif.

Le théorème de Fourier fournit le raisonnement de base. La technique consiste en l'utilisation sophistiquée de filtrages électroniques. Le théorème de Fourier affirme que n'importe quelle forme ondulatoire périodique peut être divisée en une série d'ondes sinusoïdales de fréquences différentes, d'amplitudes variables et de rapports de phases divers. Les fréquences de ces composantes d'ondes sont toutes des multiples intégrales de la répétition du fondamental de l'onde de base et sont appelées des harmoniques. La méthode la plus habituelle consiste à filtrer le signal de parole avec des filtres passe-bandes. Un filtre passe-bande est celui qui transmet des fréquences à l'intérieur d'une étendue restreinte et neutralise des fréquences aiguës et basses. La largeur du filtre détermine l'ampleur de l'étendue des fréquences qu'on laissera passer. Les fréquences définissent les limites supérieures et inférieures du passe-bande. Ces limites sont appelées les fréquences de coupage du filtre (cutoff). Les filtres qui ont un passe-bande étroit peuvent résoudre des fréquences étroitement espacées mais leurs réponses sont lourdes et lentes. Des filtres à bandes larges répondent rapidement, leur temps de résolution est très court, mais la résolution de la fréquence est pauvre. Il y a toujours ainsi un compromis entre le temps et les fréquences. La résolution de l'un ou de l'autre peut être optimisée mais pas les deux. Cette résolution est importante, car elle signifie que l'utilisateur doit sélectionner des filtres afin de montrer ou des points fins au changement temporel en utilisant des passe-bandes larges, ou les fréquences précises des composantes du signal en utilisant des filtres à bandes étroites.

Le processus utilisé pour déterminer les composantes d'une onde est appelé analyse de Fourier. C'est une technique de division par laquelle une vibration complexe est divisée dans ses composantes de vibrations harmoniques simples.

Les algorithmes des ordinateurs demandent un temps très long de résolution. Pour pallier cette contrainte, l'analyse de Fourier à transformation rapide a simplifié et accéléré le processus de résolution. Le spectre est représenté par un graphe, l'axe horizontale affiche les fréquences, l'axe verticale affiche les amplitudes. Les ondes a-périodiques affichent un spectre continu.

Comme le spectre acoustique de la parole est le produit tant de la source vocale que des caractéristiques du filtrage du tractus vocal, l'analyse du spectre nous offre la possibilité d'observer tant le fonctionnement laryngé que les mouvements articulatoires. Les méthodes d'évaluation se basent d'abord sur les **spectrogrammes** et sur une autre forme d'analyse : **l'analyse spectrale à long terme**. Cette méthode consiste à prendre la moyenne de tous les sons constituant le spectre d'un échantillon relativement long. Des investigations nous montrent que ce type d'analyse est fort prometteur en tant qu'indice de la qualité vocale. Les quelques techniques discutées se révèlent utiles dans l'évaluation de la pathologie des cordes vocales.

La recherche sur les méthodes **spectrographiques** utilisées pour analyser la qualité vocale s'est concentrée spécialement sur la caractérisation et la quantification de l'énergie acoustique située entre les fréquences harmoniques. Le contenu harmonique du signal vocal est le résultat d'une interruption quasi périodique du flux d'air généré par les cordes vocales. L'énergie non harmonique (bruit) est la corrélation acoustique d'un flux transglottique turbulent ininterrompu. Les différents aspects du bruit de la source glottique peuvent être observés dans les spectrogrammes larges et étroits (Cornut (1991)^[46]) où sont comparées les productions des sons normaux et des sons de voix enrouées. L'anormalité peut être observée dans l'espacement irrégulier des striations glottiques et dans l'environnement du bruit, évident dans l'analyse à bande large. Les différences entre les voix normales et enrouées sont plus apparentes dans les spectrogrammes à bande étroite. Cet effet est dû aux lignes claires des harmoniques vocaux.

Yanagihara et al (1967)^[356] ont évalué des orateurs hommes et femmes ayant des degrés variables d'enrouement. La spectrographie à bande étroite des voyelles

soutenues [u], [o], [a], [e], [i] à une intensité et une hauteur moyennes a pu être catégorisée en 4 types qui corrélait bien avec le degré de sévérité de l'enrouement. Les catégories de cette spectrographie sont classées selon les critères suivants :

Type 1 : les constituants harmoniques réguliers sont mêlés aux constituants de bruit, principalement dans la zone des formants des voyelles.

Type 2 : les constituants de bruit dans les formants des voyelles [e] et [i] dominent les constituants harmoniques. Des constituants de bruit léger apparaissent dans la zone des hautes fréquences au-dessus de 3000 Hz dans les voyelle [e] et [i].

Type 3 : le deuxième formant de [e] et [i] est totalement remplacé par du bruit. Celui-ci s'ajoute au-dessus de 3000 Hz. L'énergie du bruit est intensifiée.

Type 4 : les deuxièmes formants de [a], [e] et [i] sont remplacés par des composantes de bruit. Les premiers formants de ces voyelles perdent souvent leur constituant périodique au profit du bruit. Il y a davantage de constituants bruits intenses à hautes fréquences.

4.3.1.2. Les applications de l'analyse spectrale.

Une autre forme d'analyse, le spectre moyenné à long terme (SMLT), s'intéresse à la totalité du spectre sonore et s'effectue sur un échantillon spectral relativement long. Harmegnies (1992)^[113] présente une revue de la littérature en matière de spectre à long terme de la parole. Il discute des différentes sources de variation du spectre à long terme, soit les caractéristiques individuelles (sexe, profession, états émotionnels,...) les langues parlées et diverses pathologies.

La différence sexuelle dans le spectre influe sur les fréquences basses. Le SMLT des sujets masculins présente une concentration d'énergie entre 100 et 140 Hz, alors que celui des femmes se situe entre 200 et 250 Hz. Le phénomène est imputé aux caractéristiques du fondamental, dont le F_0 et la dispersion fréquentielle varient en fonction du sexe. Il relève également que les spectres des voix masculines sont moins accidentés que ceux des femmes. Les fréquences hautes présentent des différences sexuelles très ténues, voire nulles.

L'utilisation du SMLT dans le domaine musical révèle le formant du chanteur décrit antérieurement par Sundberg qui souligne que c'est particulièrement dans sa modalité parlée et non chantée que la voix des chanteurs paraît subir une évolution.

Un entraînement spécifique influe le SMLT. Il en est de même pour les voix des locuteurs professionnels où une zone énergétique apparaît dans les fréquences 3000-4000 Hz.

Le stress influence le SMLT. Harmegnies note un accroissement des énergies en faveur des fréquences comprises entre 300 et 1200 Hz. Concernant la stabilité des courbes, il rapporte que le SMLT présente une bonne stabilité intra-locuteurs pour des périodes allant de quelques jours à 3 semaines. Passé ce délai une variabilité graduelle s'installe.

Harmegnies conclut également que le changement de langue diminue la probabilité de reconnaissance correcte du SMLT.

Il confirme l'influence de la pathologie vocale et de l'altération de la fonction auditive sur le spectre.

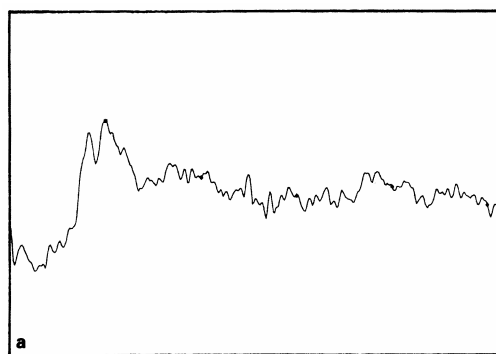
Hammarberg et al (1980)^[108] rapportent la relation entre les facteurs perceptuels et les données acoustiques spectrales extraites du spectre moyenné à long terme (SMLT). Ils analysent la voix de 17 sujets atteints de diverses pathologies. La tâche consiste à lire une histoire courte de plus ou moins 40 secondes. Quatorze juges évaluent les qualités acoustiques selon les 28 critères les plus fréquemment utilisés en clinique vocale. Une analyse factorielle met en évidence 5 axes principaux : "Instable versus stable", "soufflé versus très tendu", "hyperfonctionnel versus hypo-fonctionnel", "rude versus léger", "registre de tête versus registre de poitrine".

Les valeurs maximales des SMLT des échantillons sont retenus dans les bandes de fréquence 0-2000 Hz, 2000-5000 Hz, 5000-8000 Hz. Le facteur soufflé est corrélé avec le SMLT. Dans le cas des voix soufflées, l'énergie spectrale chute rapidement de la bande 0-2000 Hz à la bande 2000-5000 Hz où elle atteint à peu près un niveau équivalent à celui de la bande 5000-8000 Hz. Par contre, dans le cas de voix très tendues, c'est entre la bande 2000-5000 Hz et la bande 5000-8000 Hz que la baisse d'énergie est importante. Le facteur hyper/hypo-fonctionnement est aussi corrélé au SMLT. Les voix hyper-fonctionnelles sont caractérisées par une énergie importante dans les 3 bandes de fréquence, mais surtout de 2000 à 5000Hz. Les voix hypo-fonctionnelles présentent les mêmes caractéristiques que les voix soufflées. Le facteur "voix de tête / voix de poitrine" semble corrélé avec une perte d'énergie dans la bande 5000-8000 Hz. Les autres facteurs agissent de façon indépendante. Les auteurs

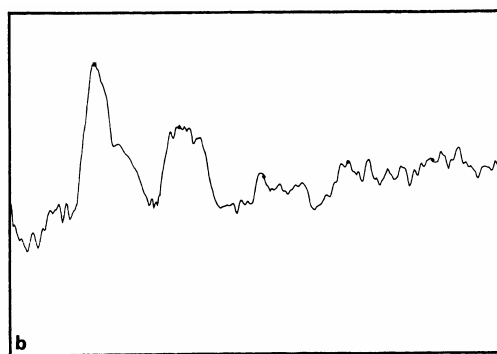
pensent dès lors que la distribution de l'énergie dans le spectre est un facteur perceptif discriminant pour les types vocaux.

Remacle et al (1989)^[249] orientent leurs analyses fréquentielles sur 150 patients dysphoniques. Chaque sujet émet un [a] tenu. L'étendue fréquentielle étudiée va de 0 à 1000 Hz. Ils obtiennent ainsi la fréquence fondamentale et les deux premiers harmoniques correspondant aux résonateurs buccal et pharyngal. Les courbes pré et post-rééducation montrent une amélioration des critères tels que l'intensité, un affinement des pics fréquents, une diminution du souffle et dès lors une meilleure sonorisation accompagnée d'une augmentation du fondamental. Remacle et al (1991)^[252] poursuivent les investigations sur des sujets présentant des nodules vocaux. Dans 75 % des cas les courbes sont modifiées ; ils observent une baisse d'intensité, un léger élargissement des pics, des harmoniques avec quelques crochets. La fréquence fondamentale par contre reste inchangée. Remacle et al (1991) constatent que ce mode d'analyse appliqué après une chirurgie partielle du larynx peut contribuer au suivi des patients et à l'objectivation de leurs qualités vocales. Ils ont observé que la réhabilitation des sujets dépendait de l'importance de la résection chirurgicale au plan glottique, de l'intégrité d'au moins une corde vocale, du traitement logopédique et de la motivation du patient. L'analyseur montrait alors une récupération vocale traduite par une amélioration de la structure acoustique. Remacle et al (1989)^[248] proposent en vue de standardiser pour l'interprétation des tracés spectraux, d'évaluer qualitativement les tracés fournis par l'analyse spectrale au moyen d'une échelle à 5 niveaux :

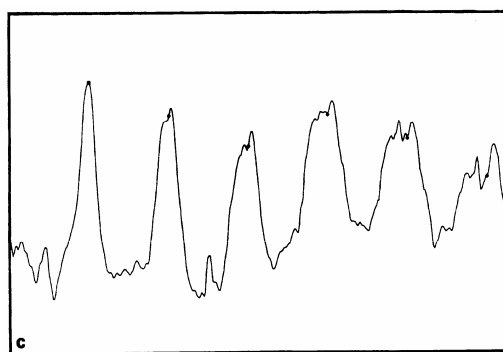
- **Niveau 0**, le tracé est plat, absence d'harmoniques et de sonorisation.
- **Niveau 1**, présence de bruit et absence de son.
- **Niveau 2**, présence de la fréquence fondamentale, pauvreté en harmoniques et présence de bruit.
- **Niveau 3**, le fondamental est présent ainsi que les harmoniques. Il y a néanmoins persistance de bruit.
- **Niveau 4**, présence du fondamental et des harmoniques. Aucun bruit n'apparaît. Le tracé est considéré comme normal [Figure 22].



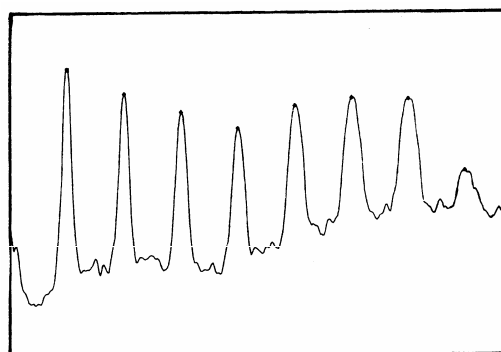
A. Tracé niveau 1



B. Tracé niveau 2



C. Tracé niveau 3



D. Tracé niveau 4

Figure 22 : Schémas d'analyse spectrale illustrant 4 niveaux de l'échelle (A,B,C,D).

Decoster et al (1997)^[56] étudient les différentes caractéristiques spectrales des voyelles [a] et [i]. Ils examinent également la mise en vibration ou "voice onset" des syllabes [pa] et [ka] en fonction du vieillissement. Il réalise une étude longitudinale sur 265 sujets et une étude cross-section sur 20 sujets incluant 5 tranches d'âge de 10 ans à partir de 20 ans. Ses résultats montrent qu'à partir de 60 ans, la différence d'amplitude entre le premier et le deuxième harmonique diminue pour les deux voyelles chez les femmes. La différence calculée entre les bandes fréquentielles en-dessous de 1000 Hz et 4000-5000 Hz augmente pour la voyelle [i] tant chez les hommes que chez les femmes.

Nawaka et al (1997)^[212] confirment par son étude la présence de "l'actor's formant" sur une combinaison de 7 voyelles. Ce formant se caractérise par une enveloppe spectrale dans la zone 3150-3700 Hz. Leino (1993)^[180], lui, trouve un pic aux alentours de 3500 Hz. La raideur de la pente spectrale s'atténue quand la sonorité ou l'intensité de la voix augmente. Ce formant a fait l'objet de peu de recherche. Le son distinctif de ce formant a été décrit acoustiquement par l'analyse spectrale à long terme par XYZ. Analogue au formant du chanteur, il est indépendant de la voyelle. Les auteurs interprètent le formant du locuteur comme une caractéristique résultant de la source glottique et des propriétés de résonance du tractus vocal.

Doskov et al (1995)^[68] examinent la présence du formant du chanteur chez 10 chanteurs d'opéra, 8 chanteurs folk, 53 non-chanteurs lors de l'émission de la voyelle [a]. Ils calculent le pourcentage d'énergie de la bande de fréquence du formant et la stabilité du fondamental. Ils constatent une grande stabilité du fondamental (>99%) chez tous les chanteurs. Par contre, leur structure formantique diffère considérablement. La concentration d'énergie dans la bande de fréquence est de 23% pour les chanteurs d'opéra, 16% pour les chanteurs folk, 7% pour les chanteurs pop, et inférieure à 4% pour les non-chanteurs.

Sundberg et al (1997)^[312] mesurent l'impact de l'ouverture de la mâchoire sur l'analyse spectrale. Dix chanteurs de différentes tessitures émettent différentes voyelles sur une échelle de 2 octaves ascendantes. Pour les voyelles [a] et [an], les chanteurs élargissent la mâchoire lorsque le fondamental approchait le premier formant à une hauteur tonale basse. Pour les autres voyelles, l'ouverture de la mâchoire est élargie à une hauteur tonale plus élevée. Cela suppose que pour ces voyelles les chanteurs emploient un autre moyen d'articulation pour augmenter le

formant 1. Les auteurs suggèrent une réduction de la constriction linguale. Le but probable de cette stratégie articulatoire est de conserver le premier formant à une hauteur tonale plus élevée que le fondamental. Cela augmente le niveau de sortie sonore sans dépenser d'effort vocal. Thooris^[in 312] affirme d'ailleurs que la qualité d'un son est inversement proportionnel à l'effort qu'il dépense.

Omori et al (1996)^[221] essayent de fournir un nouveau paramètre à l'analyse spectrale pour l'évaluation quantitative de la voix chantée. Ils mesurent le SPP qui représente les pics harmoniques les plus grand compris entre 2000 et 4000 Hz, le SPR soit le pourcentage de puissance du SPP et le pic le plus grand entre 0 et 2000 Hz. Ils rapportent que le niveau des chanteurs montre une énergie relativement plus importante dans la zone 2700 et 3400 que les non-chanteurs ($p < 0.01$). Le SPR chez les soprani est significativement plus bas que pour les autres types de voix ($p < 0.01$). Ils trouvent également une corrélation significative entre le SPR et le jugement perceptuel inter-juges.

L'analyse spectrale de la voix reflète seulement un aspect de la qualité vocale mais paraît plus proche de l'impression perceptuelle du signal acoustique de l'oreille humaine. La caractéristique sonore du chanteur professionnel est fortement liée au formant qui apparaît dans la zone 2300-2900 Hz.

Ekholm et al (1998)^[71] observent dans son étude sur le jugement perceptuel de 16 chanteurs l'influence des caractéristiques spectrales sur l'appréciation subjective. L'impression générale de la qualité vocale est influencée par l'interaction de paramètres objectifs tels que : la cadence moyenne, l'étendue moyenne, l'effet temporel. La force ou la faiblesse d'un paramètre acoustique provoque une diminution ou une augmentation des évaluations de tous les critères perceptuels. Les évaluations sont basses pour tous les critères si les échantillons manquent de vibrato. Elles sont hautes sur tous les critères si les échantillons comportent une énergie haute dans la zone du formant. La résonance corrèle avec la clarté (0.94), la couleur avec le vibrato (0.93), le critère individuel avec le score global (0.86 à 0.89). Elle observe également un lien entre la couleur et la fréquence basse du formant (-0.67), ainsi qu'entre la couleur et le formant du chanteur (0.63), et entre le timbre et le formant du chanteur. Le score est plus élevé chez les contre-ténors en voix modale que chez les falsettistes

($p=0.65$). Ils précisent qu'un score élevé pour le timbre (résonance), exige beaucoup d'énergie spectrale dans le formant.

4.3.1.3. Le rapport harmoniques/bruit.

L'intérêt de cette méthode est de pouvoir considérer isolément le bruit et le niveau des harmoniques. On a montré qu'un aspect caractéristique de l'enrouement était le remplacement de l'énergie harmonique (son périodique) par l'énergie bruit (son a-périodique). Les bases de la méthode de calcul du rapport signal/bruit est de considérer deux constituants dans la voix : des ondes parfaitement périodiques et du bruit dû au hasard.

Si on calcule l'amplitude moyenne dans chaque forme d'onde de la voyelle et que l'on considère un nombre modéré de périodes, la composante de bruit devrait se neutraliser, laissant un signal périodique pur.

Si l'onde périodique pure est envahie progressivement par du bruit dû au hasard, au fur et à mesure l'enrouement s'aggrave. Le rapport harmoniques/bruit peut être calculé par le rapport entre l'énergie acoustique du signal vocal et l'énergie acoustique du bruit. Ce rapport peut sans doute déterminer un rapport de normalité. Il est exprimé en dB. De nombreux auteurs ont essayé de quantifier ce rapport signal/ bruit. Yumoto cité par Cornut (1991)^[46] a réalisé un programme qui permet d'obtenir un quotient harmoniques/bruit faisant apparaître une valeur critique de 7.4 dB en deçà de laquelle on peut considérer la voix du sujet comme pathologique. Pour tester l'utilité de leur index, Yumoto et ses collaborateurs (1984)^[358] comparent un groupe de sujets sains ($n=42$) à un groupe de sujets dysphoniques ($n=20$) émettant un [a] soutenu. La moyenne du rapport harmoniques/bruit pour le premier groupe est de 11.9 dB ($SD=2.32$ range : 7-17) et de 1.6 dB (range : -15.2-9.6) pour le groupe pathologique. Les auteurs ont calculé la corrélation entre le taux harmonique / bruit et le type de spectrogramme (0.849), la corrélation entre les taux H/N pré et post-opératoire et le type de spectrogramme (0.944) et le taux H/N et l'échelle psychophysique G.R.B.A.S de la raucité (0.809).

Dalléas et al (1987)^[48] parle d'un rapport signal / bruit (S/B) discriminant entre voix normales (16.6 et 28.8 dB) et voix pathologiques (1.9 et 17.3 dB).

La mesure du rapport S/B varie en fonction du support employé pour l'analyse de l'échantillon. Gelfer et al (1995)^[88] analysent la voix de 30 sujets à l'aide du C Speech Program d'IBM. Via un enregistreur numérique, les valeurs S/B augmentent toutes comme celles du shimmer et du jitter mais de manière moins importante (0.741 versus 0.481). Les résultats de Titze et al^[325] et Doherty et al^[67] sont également supérieurs à ceux obtenus lors de la prise d'informations directes.

Wolfe et al (1997)^[352] étudient les corrélations entre le type de pathologie vocale et la prédiction de la gravité de la dysphonie. Cent deux échantillons sont répartis en trois catégories : voix soufflée, voix enrouée, voix rauque. Le jury évalue sur une échelle à 7 degrés la gravité de la dysphonie. Les résultats montrent que le rapport signal/bruit explique 58% de la variance perceptuelle des voix enrouées et 26 % de la variance des voix soufflées. Le rapport signal / bruit corrèle avec les voix soufflées ($r=-0.51$, $p=-0.035$), et les voix enrouées ($r=-0.76$, $p=0.033$). Ils citent de Krom et al (1995)^[55] dont les résultats sont similaires. Martin et al (1995)^[189] (-0.85), Eskenazi et al (1990)^[77], Kojima et al (1980)^[167] (-0.86), Yumoto et al (1984)^[358] (-0.80) corroborent également ces résultats.

Martin et al (1995)^[189] rapportent que le HNR prédit plus la raucité ($r=-0.85$) que le shimmer (0.66) et le jitter (0.50).

Hirano et al (1986)^[124] trouvent que le facteur (G) corrèlent négativement avec le rapport signal / bruit. La raucité tend à être associée avec une étendue fréquentielle réduite. La sensation de faiblesse corrèle avec une grande fluctuation en amplitude et un haut niveau de bruit. Les patients dont la qualité vocale est caractérisée par une sensation de faiblesse ont une étendue fréquentielle et une étendue en amplitude réduites. Le paramètre de sensation de forçage (S) ne corrèle avec aucun paramètre acoustique.

L'analyse spectrale a également suscité l'intérêt de beaucoup de chercheurs dans le domaine de la voix chantée (Morris et al (1995)^[203] ; Miller (1987)^[200] ; Miller et al (1990)^[198], et al (1993)^[199] ; Novak et al (1995)^[217] ; Nawka et al (1997)^[212] ; Omori et al (1996)^[221] ; Siegwart et al (1995)^[283] ; Sundberg (1970)^[303], et al (1993)^[309] ; Griffin et al (1995)^[104] ; Harmegnies (1988)^[112], (1992)^[113] ; Blaylock (1999)^[23] ; Dostkov et al (1995)^[68] ; Ekholm et al (1998)^[71]) tant dans un premier temps pour l'identification

et la description du formant du chanteur que la précision des registres, la corrélation avec le degré d'aperture buccale, l'influence des mouvements du voile du palais, de la langue, la configuration du vibrato et l'impact de différentes techniques vocales sur l'analyse spectrale. Certaines de ces études ont été décrites au point 2.2.2.3.

4.3.2. Le "Multi Dimensional Voice Program" (M.D.V.P.).



Figure 23 : Illustration du Computer Speech Lab de Kay Elemetrics.

Le Computer Speech Lab (C.S.L.)^[159] est conçu pour l'analyse acoustique d'échantillons vocaux, basée sur différents algorithmes mathématiques. Les analyses peuvent se faire sur des échantillons de paroles (phrases que l'on peut éventuellement comparer et juxtaposer). Elles peuvent également être réalisées sur une voyelle soutenue. Le programme vocal multi-dimensionnel (M.D.V.P.) est une option du C.S.L.. C'est une évaluation quantitative clinique de la qualité vocale calculée sur 33 paramètres à partir d'une phonation soutenue simple, telle une voyelle.

Les mesures acoustiques sont les suivantes :

Ce programme est conçu pour l'analyse de la voix parlée. Il enregistre 33 paramètres de la voix à partir d'une simple vocalisation. Ce programme utilise le système Computerized Speech Lab (C.S.L) pour capter le signal, l'analyser et le restituer auditivement. Le C.S.L utilisé pour cette étude était un modèle 4300 B Software Version 5.X (Kay Elemetrics Corp., Pine Brook, New-Jersey, USA). Les 33 paramètres vocaux analysés sont retranscrits sous forme de données chiffrées. Est adjointe une conversion graphique du signal de ces paramètres, qui donne une représentation visuelle de la vocalisation du sujet. Les données graphiques et chiffrées peuvent être imprimées. En cas de représentation graphique, le M.D.V.P. affiche plusieurs fenêtres numérotées de A à G [annexe 7].

4.3.2.1. La description des paramètres.

a) Les mesures de la fréquence fondamentale :

- **Fo** : la fréquence fondamentale moyenne exprimée en Hz.
- **Flo** : la fréquence fondamentale la plus basse pour toutes les périodes extraites.
- **FHI** : la fréquence fondamentale la plus haute pour toutes les périodes extraites.
- **To** : la période moyenne de la hauteur tonale pour toutes les périodes de hauteur tonale.
- **STD** : la déviation standard de la fréquence fondamentale dans l'échantillon vocal analysé.
- **PFR** : l'étendue de la fréquence fondamentale phonatoire entre FHI et Flo (1/2 ton) ou "fréquence range".
- **Tsam** : la longueur de l'échantillon des données analysées.
- **PER** : le nombre de périodes de la hauteur tonale détecté dans l'échantillon vocale.
- **SEG** : le nombre total de segments calculés par l'ordinateur durant l'analyse d'auto-corrélation.

b) Les mesures de la perturbation fréquentielle à court et à long terme.

- **Jitta** : ou jitter absolu, donne une évaluation de la variation période à période de la hauteur tonale dans l'échantillon vocal analysé.
- **Jitter** : le jitter en pourcentage donne une évaluation de la variation de la hauteur de l'échantillon vocal enregistré. Il représente la variation cycle à cycle à très court terme.
- **PPQ** : le quotient de la perturbation de la période de la hauteur tonale donne une évaluation de la variabilité de la hauteur tonale dans l'échantillon vocal analysé en régularisant 5 périodes du facteur.
- **RAP** : la perturbation moyenne relative donne une évaluation de la variabilité de la période de la hauteur tonale dans les échantillons analysés en régularisant 3 périodes du facteur.

- **SPPQ** : le quotient de perturbation de la période du pic régulier donne une évaluation de la variabilité à court terme et à long terme de la période du pic dans l'échantillon vocal analysé.
- **VFo** : la variation de la fréquence fondamentale représente la déviation standard relative de période à période de la fondamentale. Cette mesure reflète la variation à très long terme de Fo pour tous les échantillons vocaux analysés.

c) Mesures de la perturbation en amplitude à court et à long terme.

- **SHDB** : le shimmer est une autre composante de ce qui est perçu comme une raucité. Le shimmer est mieux mesuré durant la phonation soutenue d'une voyelle. Le shimmer en dB donne une évaluation de la variabilité d'amplitude cycle à cycle dans l'échantillon vocal analysé.
- **SHIM %** : le shimmer en pour-cent donne une évaluation de la variabilité d'amplitude cycle à cycle de l'échantillon vocal analysé. Cette mesure représente la variabilité relative période à période (à très court terme) de l'amplitude pic à pic.
- **APQ** : le rapport de la perturbation de l'amplitude
- **sAPQ** : le rapport de la perturbation de l'amplitude régularisée, elle donne une évaluation à long ou à court terme de l'amplitude pic à pic dans l'échantillon vocal analysé. Le facteur de régularité peut être défini par l'utilisateur.
- **vAM** : représente la déviation standard relative de l'amplitude cycle à cycle.

d) Mesures des ruptures vocales.

- **DVB** : calcule le degré de cassures vocales en pourcentage.
- **NVB** : nombre de cassures vocales montrant combien de fois le fondamental généré est interrompu depuis le début jusqu'à la fin de la dernière zone voisée.

e) Mesures des composantes sub harmoniques.

- **DSH** : calcule le degré de composantes sub-harmoniques en pour-cent.
- **NHS** : nombre de segments trouvés sans harmonique durant l'analyse.

f) Mesures liées à l'irrégularité vocale.

- **DUV** : un degré de moindre voisement estimé par une évaluation relative de zones sans harmonique (dans laquelle F_0 ne peut être détecté) dans un échantillon vocal.
- **NUV** : nombre de segments non voisés détectés durant l'analyse d'auto corrélation.

g) Mesures liées au bruit.

- **NHR** : Le rapport harmoniques/bruit compare la période ou le signal régulier produit par les cordes vocales, avec le bruit ou le signal apériodique des cordes vocales et du tractus vocal. La voix la plus rauque ou la plus bruyante a un rapport harmoniques/bruit plus bas. La voix claire a un rapport élevé. Les algorithmes pour calculer NHR sont inclus dans le programme.
- **VTI** : l'index de turbulence vocale est un rapport moyen de l'énergie fréquentielle haute du spectre non harmonique dans l'étendue 2800-5800 Hz à l'énergie spectrale non harmonique s'étendant de 70-4500 Hz dans la zone du signal où l'influence de la variation de fréquence et d'amplitude, les cassures vocales et les composantes sous harmoniques sont minimales. VTI représente le niveau d'énergie relatif au bruit fréquentiel haut. C'est grandement corrélé avec la turbulence causée par une adduction incomplète.
- **SPI** : ce paramètre n'est pas une mesure du bruit mais plutôt la structure harmonique de l'index de phonation du spectre (souple, doux,...). C'est un rapport moyen de l'énergie de la fréquence la plus basse (70-1600 Hz) à l'énergie harmonique de la fréquence la plus haute (1600-4500 Hz). La valeur augmentée de SPI peut être une indication de l'adduction incomplète ou approximative des cordes vocales durant la phonation. SPI est très sensible à la structure du formant de la voyelle, parce que la voyelle avec une énergie fréquentielle haute diminuée donne un SPI plus haut. Seules les valeurs calculées par l'ordinateur pour la même voyelle peuvent être comparées. La voyelle [a] est recommandée.

h) Mesures de tremblements.

- **Ftri** : l'index de "tremblement" de l'intensité en pour-cent traduit le rapport de l'ampleur fréquentielle de la fréquence basse la plus forte avec l'ampleur fréquentielle totale du signal analysé.
- **Atri** : l'index de "tremblement" de l'amplitude en pour-cent

Les conditions d'enregistrement à respecter sont de conserver le même environnement , d'employer le même microphone, de le placer à 30 cm de la bouche du sujet.

Il est important de se rappeler que les valeurs d'intensité sont directement liées à la distance du microphone aux lèvres. Quelques phrases standardisées peuvent être employées. Les phrases doivent être appropriées à l'élément analysé. Les phrases doivent être brèves pour éviter les éventuelles difficultés de rétention mémorielle, et faciliter l'analyse des résultats. Les échantillons vocaux pourraient inclure la voyelle [a] et une phrase standard de 3 sec. Le C.S.L. peut stocker jusqu'à 8 combinaisons de courbes sonores.

4.3.3. Les stations E.V.A. 1 et 2.



Figure 24 A : Dispositif E.V.A. :boîtier électronique et micro-ordinateur.

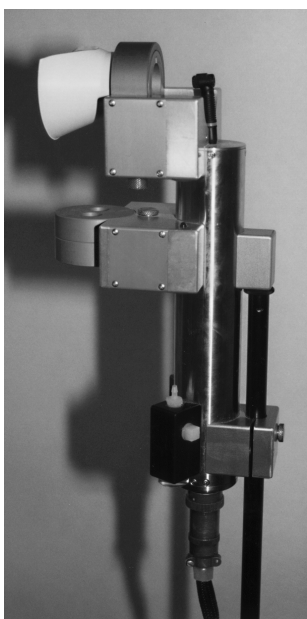


Figure 24 B : la pièce à main avec l'embouchure buccale en silicone souple.

L'appareillage E.V.A.^[293, 296] fut mis au point au début des années 90 par l'équipe de la fédération O.R.L du C.H.U. de la Timone (France) en collaboration avec le laboratoire "Parole et Langage" de l'Université de Provence. E.V.A. est un outil d'évaluation objective de la voix reposant sur l'analyse simultanée des paramètres

acoustiques et aérodynamiques d'un son tenu. Le matériel comprend un ensemble de capteurs et de circuits électroniques reliés à un micro-ordinateur au niveau duquel le logiciel E.V.A. gère les données, les calculs et leur présentation. Les capteurs acoustiques, de débit et de pression sont contenus dans une pièce à main fixée sur un pied adaptable à la taille du patient. Des embouchures buccales et/ou nasales de silicone souple reliaient hermétiquement la sphère orale du patient au système. Le microphone est placé à 2 cm de l'embouchure. Très directionnel, il autorise la réalisation des mesures dans un endroit calme. Il ne nécessite pas de cabine anéchoïde. Le dispositif électronique de parole comprenant les filtres du signal et les instruments de mesure des paramètres vocaux sont intégrés dans un boîtier. Il se compose de dix modules indépendants. Chaque module est doté d'un système de réglage adaptable aux caractéristiques de la voix du patient. La mesure de fréquence (F_0) est un mélodimètre d'une précision de 1 Hz pour un signal sinusoïdal de 1000 Hz. Le mélodimètre repose sur le principe d'une détection électronique de la période après filtrage, le "zero-crossing" tel que recommandé par Titze (1995)^[328]. Le "zero-crossing" est une méthode dite automatique. Elle consiste à filtrer le signal à analyser afin d'en éliminer les harmoniques. Ensuite, le nombre de passages par "zéro" du signal est décompté et divisé par deux. La méthode d'extraction de la période utilisée avec la station E.V.A. 1 est une méthode de détection électronique du type "zero-crossing" ascendant après filtrage passe-bas. Ils ont utilisé une bande passante de 0 à 250 Hz pour les voix masculines et de 0 à 500 Hz pour les voix féminines et enfantines (Titze (1996) ^[in 98]). Le détecteur d'intensité correspond à la mesure du logarithme de la valeur efficace du signal de parole à ± 2 %. Les mesures de débit phonatoire (DAB) exprimées en cm^3/sec sont calculées par un pneumotachographe thermostaté à grille linéaire pour des valeurs inférieures à 1 l/sec. Tous les paramètres de la voix sont digitalisés simultanément. Les mesures recueillies s'affichent sous forme de courbes d'évolution temporelle de la fréquence fondamentale (F_0), d'intensité (INT) et de débit d'air buccal (DAB) qui correspond en l'absence de fuite nasale au débit d'air glottique. La moyenne, l'écart-type, le coefficient de variation sont calculés pour chaque paramètre. Des mesures de variabilité de F_0 et de Int sont exprimées par le jitter et le shimmer en pour-cent et en valeur absolue, ainsi que le shimmer débit pour le débit d'air buccal. La mesure de la fuite glottique est également disponible. Elle consiste en un rapport du DAB sur l'INT, soit la quantité d'air nécessaire pour émettre un son de 1 décibel en 1 seconde ($\text{cm}^3/\text{dB/s}$). Le logiciel étudie également la représentation

spectrale de l'échantillon vocal enregistré. Le graphique reprend l'intensité (en dB) en fonction de la fréquence (en Hz) avec repérage des harmoniques et des points d'émergence du bruit. Le temps nécessaire pour réaliser les mesures n'excède pas 10 minutes par sujet.

La station E.V.A. 2^[296] repose sur les mêmes principes que E.V.A. 1^[293]. Le nouveau dispositif offre plus de possibilités d'analyse et facilite le réglage des paramètres. En un mot, le dispositif est plus convivial [annexe 8]. En voici les possibilités :

4.3.3.1. La description des paramètres ^[296] :

- **L'évaluation vocale ou profil vocal**

Le sujet prononce un [a] tenu. L'expérimentateur précise une zone d'analyse sur une portion de l'échantillon. A partir de cette fenêtre d'observation, le logiciel calcule des indices tels que :

- la moyenne, l'écart-type, le coefficient de variation de la fondamentale, de l'intensité et du débit d'air oral,
- l'instabilité de la fréquence de vibration représentée par le jitter moyen, le jitter factor, le jitter ratio, le "Relative Average Perturbation" (RAP),
- l'instabilité de la vibration en amplitude : shimmer, "Amplitude Perturbation Quotient" (APQ),
- le rapport signal/bruit,
- l'indice de fuite glottique.

a) Mesures de la fréquence.

- **F₀** (Hz) : la fondamentale moyenne apporte une mesure globale de la hauteur.

- **STD dev** (Hz) : l'instabilité à moyen terme du fondamental se traduit par des variations de fréquence dans le temps. Elle se mesure par l'écart-type du fondamental qui correspond à l'ampleur en Hz des variations de F_0 .
- **C.V.(%)** : le coefficient de variation permet de relativiser l'écart-type en le comparant au F_0 moyen. Il correspond donc à l'ampleur en pourcentage des variations de F_0 par rapport au F_0 moyen. Par exemple, à un écart-type de 4.9 Hz pour un F_0 moyen de 180 Hz correspond un coefficient de variation de 2.7%. un écart-type de 4.9 Hz pour un F_0 moyen de 500 Hz correspond à un coefficient de variation de 0.98%. Cette valeur est plus normale que la précédente. Les auteurs considèrent donc que le coefficient de variation du fondamental est le meilleur indice pour explorer la stabilité de la fréquence fondamentale à moyen terme. Il devient important pour des phénomènes de vibrato pathologique, de chevrottement, d'instabilité d'origine neurologique.
- **Jitt ou CTC perturbation** (Hz) : le jitter moyen permet de calculer l'instabilité à court terme du F_0 . Elle se traduit par des variations de fréquence entre chaque cycle d'oscillation. Cette mesure est la moyenne de la différence du F_0 entre deux cycles de vibration consécutifs.
- **RAP** : la perturbation moyenne relative mesure la moyenne des variations de trois périodes consécutives rapportées à la période moyenne du signal observé.
- **Jitter Factor** (%) : le jitter factor permet de relativiser le jitter moyen en le comparant au F_0 moyen. Par exemple à un jitter moyen de 0.677 HZ pour un F_0 moyen de 180 HZ correspond un jitter factor de 0.38%. Le même jitter moyen pour un F_0 de 80 HZ donne un jitter factor de 0.84%. Le jitter factor est donc un meilleur indice pour explorer la stabilité de F_0 à court terme.
- **Jitter ratio** (%) : le jitter ratio mesure la moyenne des variations de période entre deux cycles de vibration consécutifs et relativise cette valeur par la période moyenne du signal observé.

b) Mesures de l'intensité :

- **INT moyen** (dB) : l'intensité moyenne apporte une mesure globale de la force sonore de la voix.
- **INT Std dev** (dB) : l'écart-type de l'intensité correspond à l'ampleur en décibels des variations d'intensité autour de la moyenne.

- **INT C.V. (%)** : le coefficient de variation permet de relativiser l'écart-type en le comparant à l'intensité moyenne. Il correspond donc à l'ampleur en pourcentage des variations d'intensité par rapport à l'intensité moyenne.
- **Shim moyen** : le shimmer moyen représente l'instabilité à court terme de l'amplitude de la vibration. Il se traduit par des variations d'amplitude entre chaque cycle d'oscillation. Cette mesure est la moyenne des rapports d'amplitude entre deux cycles de vibration consécutifs. Ces variations sont calculées cycle à cycle de manière très précise.
- **APQ** : le quotient de perturbation moyenne de l'intensité mesure la moyenne des variations d'amplitude en tenant compte de 11 périodes consécutives, le tout rapporté à l'amplitude moyenne du signal observé.
- **Shim factor (%)** : le shimmer factor permet de relativiser le shimmer moyen en le comparant à l'amplitude moyenne.

c) Mesures du rapport signal/bruit (extraites de l'analyse spectrale.).

- **Sr** : ou "Relative Signal Intensity" correspond à l'énergie en pourcentage du fondamental et des harmoniques comparée à l'énergie totale.
- **H/N** : ou "Harmonic to Noise Ratio" correspond à l'énergie du fondamental et de ses harmoniques comparée à l'énergie du bruit.
- **NNE** : ou "Normalised Noise Energy" correspond à l'énergie du bruit comparée à l'énergie totale (en dB).

Sr, H/N et NNE sont calculées sur tout le spectre.

Les valeurs de Sr hf, H/N hf, NNE hf sont calculées sur la partie du spectre au-dessus de 1000 Hz.

d) Mesure aérodynamique :

- l'indice de **fuite glottique** (cm³/s/db) explore la présence ou l'absence de fuite d'air à travers la glotte. Cette mesure est égale au DAB moyen rapporté à l'intensité moyenne.

- **Le temps maximal phonatoire.**

Ce programme est destiné à évaluer le rendement laryngé, soit le rapport entre la quantité d'énergie aérienne fournie et la durée de l'émission vocale rendant compte indirectement de la fuite glottique.

- **L'efficacité glottique ou Pression sous-glottique estimée (PSGE).**

- **L'efficacité glottique** (dB/hPa) : évalue le rendement glottique, soit le rapport entre la quantité d'énergie acoustique émise et la quantité d'énergie aérodynamique nécessaire à cette émission. Ces indices explorent la notion de forçage vocal. Le principe est basé sur la méthode de Smitheran et Hixon (1981)^[286]. Le sujet émet une série de [pa pa pa].
- **La PSGE** (hPa) : la pression sous-glottique estimée est mesurée juste avant l'explosion du [p] sur la courbe de pression. Nous détaillerons cette mesure au point 4.2.3.2.
- **La Résistance glottique** (hPa/dm³/s) : est le rapport de la pression sur le DAB.

- **La spirométrie**

La **spirométrie** permet d'évaluer les volumes expiratoires et de détecter d'éventuelles insuffisances respiratoires.

- **La nasalité**

La mesure de la **fuite nasale** est destinée à évaluer le degré d'insuffisance vélaire congénitale ou acquise.

- **Le phonétoqramme**

Le **phonétoqramme** évalue l'étendue vocale d'un sujet en enregistrant l'intensité minimale et maximale en fonction de la hauteur de la voix.

Il mesure également la capacité vitale, l'analyse prosodique, et la voix chantée.

Giovanni et al (1996)^[96] précisent qu' E.V.A. n'est pas un modèle pour la voix en co-articulation. Il suggère le choix de la voyelle [a] dont le spectre, riche en harmoniques graves et aigus, rend sensible aux moindres altérations vocales. Le [a] est aussi la voyelle la plus ouverte et par conséquent la moins affectée par les obstacles du conduit vocal supra-glottique. Le [a] entraîne peu de turbulences aériennes et la totalité de l'air expiratoire passe par la bouche, ce qui simplifie la mesure des débits et la rend plus fiable. De plus, elle est efficace au dépistage d'une insuffisance vélaire. Elle est simple à produire, demande peu d'investissement de la part du sujet, facilement reproductible et moins perturbée par l'émotion que la lecture d'un texte. Le testing en est facilité et rendu plus objectif.

4.3.3.2. Les mesures de la pression sous-glottique.

La pression sous-glottique estimée a été intégrée à la station E.V.A. 2 Dans ce paragraphe, nous discutons de l'intérêt croissant pour cette mesure dans le domaine de la recherche.

L'émission sonore est le résultat d'une utilisation soignée et précise de la pression d'air produite par le système respiratoire. Pendant la phonation, l'accolement des cordes vocales crée un obstacle. Celui-ci entraîne une élévation de la pression en amont de la glotte, la pression sous-glottique. Cette élévation de pression va donner l'impulsion nécessaire à la vibration des cordes vocales. La pression sous-glottique joue donc un rôle déterminant dans la vibration des cordes vocales. Il est utile de connaître la pression d'air dans une zone donnée du tractus vocal et de la comparer à la pression dans d'autres zones données. Le changement des valeurs de pression qui résulte de l'activité parlée permet de comprendre la structure et le fonctionnement du

tractus vocal et aussi au clinicien de déduire un certain nombre d'informations concernant la nature des troubles laryngés.

Plusieurs méthodes pour mesurer la pression sous-glottique (Giovanni 1998^[98], Jiang et al (1999)^[154]) ont été décrites :

- Directement en utilisant un **capteur trachéal trans-laryngé**. Celui-ci est placé dans la région sous-glottique à l'aide d'une aiguille, à travers la membrane crico-thyroïdienne ou entre les cordes vocales (ponction). Netsell (1969)^[214] a utilisé une aiguille de 4 cm de longueur avec un diamètre inférieur à 0.055 cm insérée entre le premier et le deuxième anneau trachéal. L'aiguille est attachée à un transducteur de pression par un tube en polyéthylène. Les désavantages de la méthode néanmoins sont évidents et significatifs. Même si la procédure n'est pas spécialement douloureuse, elle peut provoquer une anxiété considérable chez le patient. Du point de vue technique, le placement du tube doit être fait par un praticien compétent. Des risques d'endommagement ou de saignement existent. Ces difficultés ont activé le développement d'approches moins invasives.
- Directement, en utilisant un **transducteur¹⁹ de pression miniature inséré par voie nasale** jusqu'à ce que son extrémité atteigne le dessous du voile du palais. Le tube est donc situé dans l'oropharynx. L'extrémité du tube est dirigé ainsi vers le flux d'air. L'avantage de ce placement nasal est qu'il n'y a pas d'interférence avec les mouvements articulatoires, pas d'activité linguale ou labiale vive qui déplacerait le tube. Le mouvement vélaire ne semble pas compromettre l'évaluation de manière significative. Le tube sera néanmoins choisi assez fin (2 mm). Une seringue hypodermique est connectée à une jauge de pression munie d'un robinet à trois voies. Un jet d'air envoyé dans le tube évite que la salive ne l'emplisse.
- Indirectement, en utilisant un **ballonnet œsophagien**. La PSG est ainsi calculée à partir de la pression œsophagienne quand le volume pulmonaire est constant.

¹⁹ Un **transducteur** est un dispositif qui reçoit de l'énergie d'un ou plusieurs systèmes ou milieux et qui fournit une énergie correspondante mais d'une autre nature. Le transducteur électro-acoustique transforme l'énergie électrique en énergie acoustique et vice versa.

- Indirectement, en plaçant le sujet dans un **pléthysmographe**²⁰.

Plusieurs études ont présenté des résultats de la mesure de la PSG chez des sujets normaux. Le problème du calibrage rend cependant la comparaison des valeurs difficile, et ce, sans tenir compte de la variété des méthodes de mesure.

Tableau 40 : Valeurs de pression sous-glottique en fonction de 3 méthodes. Auteurs cités par Giovanni (1998)^[98].

Auteurs	Années	Méthodes	PSG (hpa)
Schutte ^[279]	1980	Ballonnet oesophagien	13.9
Iwata ^[150]	1988	Capteur trachéal trans-laryngé	10.2
Plant et al ^[235]	1998	Ponction	8 à 16

Ces différentes méthodes et particulièrement la ponction restent invasives, inconfortables pour le sujet et difficiles à envisager en pratique clinique.

Smitheran et Hixon (1981)^[286] ont mis au point une méthode non invasive pour estimer la PSG de façon indirecte : "*The airway interrupted method*".

Lors d'une succession rapide de [pi] [pi] [pi], il est possible de considérer comme constante la PSG générée par la pression pulmonaire phonatoire. Grâce à un capteur de pression placé dans la bouche, la pression intraorale (PIO) peut être mesurée.

Sur l'émission du [i], les lèvres sont ouvertes, les cordes vocales sont en vibration et la PIO est égale à la pression atmosphérique.

Par contre, sur la tenue du [p], les lèvres sont fermées, la glotte est ouverte et les cordes vocales ne vibrent pas. Un équilibre des pressions s'établit dans le conduit vocal : la PSG peut être estimée en mesurant la PIO [figure 25] (Giovanni (1998)^[98], Hertegard et al (1995)^[115], Zajac (1998)^[360], SQLab^[296]).

²⁰ Le **Pléthysmographe** est une enceinte close utilisée en exploration fonctionnelle respiratoire pour mesurer certains paramètres, en particulier les volumes pulmonaires et les résistances bronchiques. On distingue le pléthysmographe volumétrique, à pression constante où l'on enregistre les variations de volume et le pléthysmographe barométrique, à volume constant où l'on enregistre les variations de pression ^[15].

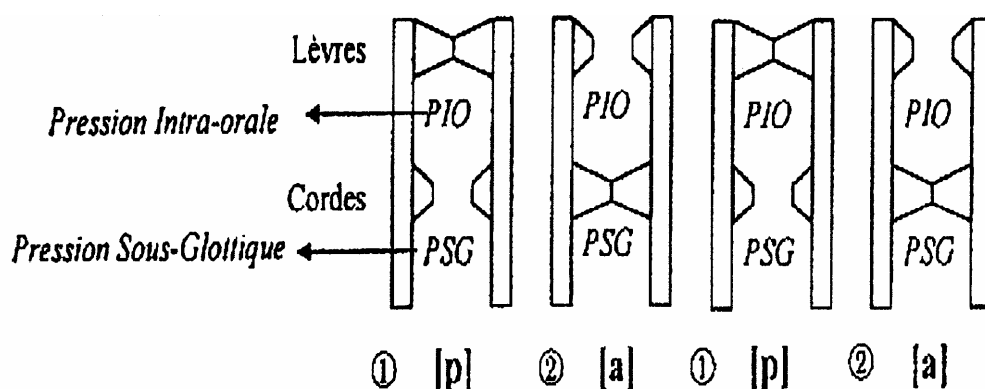


Figure 25 : Principe de l'estimation de la PSG par l'intermédiaire de la PIO, SQ lab. ^[296]

Ainsi, lors d'une succession rapide de [pi] [pi] [pi], si la PSG est supposée constante, la PIO mesurée juste avant l'explosion du [p] permet une estimation satisfaisante de la PSG.

Elle est alors appelée "**pression sous-glottique estimée**" (**PSGE**).

Plusieurs études ont démontré la fiabilité de cette méthode grâce à des mesures simultanées trans-trachéales et intra-orales.

Löfquist et al (1982)^[186] ont effectué des mesures de la pression au-dessus et en-dessous de la glotte grâce à un transducteur de pression miniature. Le coefficient de corrélation entre la PSG et la PIO était de 0.92.

De même Hertegard et al (1995)^[115], ont mesuré la PIO et la PSG par ponction trachéale et ont enregistré le signal électroglottographique pendant la phonation d'un orateur ne présentant aucune pathologie. La PSG a aussi été estimée à partir de la PIO juste avant l'explosion d'un [p]. Les résultats montrent que les valeurs des pressions diffèrent très peu (en moyenne valeurs inférieures à 2%) et que les mesures des pressions sont hautement corrélées ($r = 0.98$).

Une autre étude a été réalisée par McHenry et al (1995)^[195], les auteurs ont procédé à des mesures directes du flux d'air, des pressions sous-glottique, intraorale

et pharyngée chez quatre hommes et quatre femmes ne présentant aucune pathologie vocale.

La procédure était particulièrement invasive :

- La pression orale était mesurée à l'aide d'un cathéter de polyéthylène placé à l'intérieur de la bouche et entre les lèvres.
- La pression pharyngale était mesurée à l'aide d'un cathéter passé à travers les fosses nasales et positionné au-dessus de l'épiglotte.
- La pression sous-glottique était mesurée à l'aide d'un cathéter placé dans la trachée après une anesthésie sous-cutanée [figure 26].

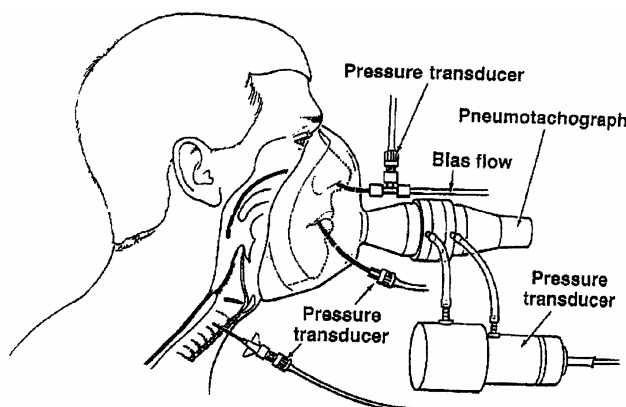


Figure 26 : Schéma de la procédure de McHenry et al.^[195].

Les sujets devaient produire des suites de syllabes comme [pi] [pi] [pi] par exemple, de façon normale, à voix haute, à voix douce, à voix soufflée et forte. L'objectif des auteurs était de vérifier une éventuelle différence entre les mesures directes et indirectes de la résistance²¹ du conduit laryngé selon le type de syllabe ou la condition vocale.

²¹ *La résistance du conduit laryngé est calculée en divisant le pic de pression intra-orale précédant l'ouverture des lèvres par la quantité du flux d'air libéré pendant le segment voisé.*

Tableau 41 : Valeurs de pression intra-orale et transglottique (cm H₂O) rapportées par McHenry 1995^[195].

4 hommes /[pa]	Intensité	Pression Orale		Pression TG*		Flux (ml/sec)		PO-PTG**
Moyenne	Normale	5.71	0.32	6.27	0.19	118.8	3.40	-0.56
DS***		0.53	0.18	0.71	0.15	35.10	4	
Moyenne	Forte	9.67	0.35	9.95	0.35	160.50	7.60	-0.28
DS		1.01	0.19	0.55	0.25	28.30	11	
Moyenne	Douce	5.97	0.19	5.38	0.16	140.9	9.7	0.59
DS		1.27	0.10	1.25	0.09	42.80	13	
Moyenne	Soufflée	10.19	0.03	8.58	0.09	359.6	6.1	1.61
DS		1.71	0.05	1.54	0.13	119.4	6.3	
Moyenne	Serrée	13.13	1.05	14.33	0.95	223.6	9.90	-1.20
DS								
4 Femmes/[pa]								
Moyenne	Normale	5.93	0.52	6.01	0.31	111.10	5.30	-0.08
DS		1.41	0.33	1.59	0.21	40.20	4.90	
Moyenne	Forte	9.88	0.26	9.35	0.40	112.60	3.00	0.53
DS		3.25	0.15	3.05	0.25	56.30	1.80	
Moyenne	Douce	5.56	0.19	5.83	0.28	158.10	2.70	-0.27
DS		2.15	0.17	1.93	0.12	76.90	2.30	
Moyenne	Soufflée	8.62	0.31	7.49	0.36	225.90	9.80	1.13
DS		4.43	0.28	3.33	0.30	156.40	7.00	
Moyenne	Serrée	10.40	1.10	11.85	0.70	126.40	13.30	-1.45
DS		5.68	0.47	4.43	0.27	16.70	8.30	

Légende : *TG :transglottique ; PO-PTG : pression orale-pression transglottique ; *** DS : déviation standard.

La comparaison des mesures directes et indirectes a révélé une correspondance généralement satisfaisante, en particulier pour les syllabes [pi] (valeur moyenne pour les mesures directes chez les hommes = 6.46 hpa et indirectes = 5.48 hpa / valeur moyenne pour les mesures directes chez les femmes = 6.58 hpa et indirectes 6.03 hpa).

La pression sous-glottique varie directement avec l'intensité vocale. Le rapport exact n'est pas linéaire. Il diffère quantitativement entre les individus. Isshiki en 1964^[146] rapporte que la pression sous-glottique mesurée par ponction trachéale chez un de ses sujets masculins variait de 3 à 25 cm H₂O alors que le niveau de pression sonore augmentait de 65 à 95 DB.

Slis en 1970 ^[285] observe des pressions sous-glottiques à une intensité mesurée en voix chuchotée, en voix normale et lors du cri. Les valeurs sont de respectivement 10 cm H₂O ; 11.5 cm H₂O et 19 cm H₂O.

Les sujets de l'étude de McHenry et al (1995)^[195] répètent des séries de 7 explosives à intensité normale, forte, faible en simulant des voix soufflées et forcées. Il conclut que l'utilisation de la pression intra-orale pour estimer la pression transglottique est cliniquement raisonnable. Il conseille la prudence dans l'interprétation des valeurs obtenues en voix soufflée et en voix forcée. La résistance du conduit laryngé peut être sous-estimée dans le premier cas et surestimée dans le second. Les évaluations de la résistance seraient valides lorsque le flux d'air est compris entre 90 et 300 ml/sec. Au delà de 300 ml/sec, la pression transglottique est inférieure à la pression intra-orale dans 100% des cas. En deçà de 90 ml/sec, la pression translaryngée est supérieure à la pression intra-orale dans 73 % des cas. Le fait que toutes les données aient été obtenues simultanément valide les estimations indirectes de la résistance du conduit d'air.

Les auteurs concluent cependant qu'il faut rester prudent par rapport aux comportements extrêmes que sont les voix soufflées et fortes.

En effet, la résistance du conduit laryngé risque d'être sous-estimée dans le cas de la prononciation à voix forte, et surestimée dans le cas de la voix soufflée.

La meilleure estimation est fournie par la voix normale.

La pression intra-orale peut être également examinée pendant la parole. Les résultats donnés par Baken et al (1987)^[14] sont calculés sur la phrase "*the top is on the table*". La pression intra-orale sur le premier [T] était à peu près de 12 cm H₂O. Le [T] de "table" a été produit avec 10 cm H₂O. Le [P] a été émis avec +/- 9 cm H₂O. Le [B] a été émis à 5 cm H₂O. L'uniformité relative des courbes des pics de pression est cohérente avec la nature de l'émission. Elles peuvent être considérées comme étant un indicateur de la stabilité du temps de parole.

La forme du pic de pression pour les explosives est significative pour deux raisons : La chute abrupte de la pression intra-orale coïncide avec l'éclat de l'explosive et révèle un fonctionnement normal. En cas de dysfonctionnement de la parole, les pics de pression deviennent moins distincts, plus larges, plus arrondis, et se fondent les uns avec les autres.

En conclusion, toutes ces études confirment la pertinence de l'estimation de la PSG par la PIO ; néanmoins quelques exigences sont de mise (Hertegard et al (1995)^[115]) :

- 1) La résistance glottique doit être nulle et les cordes vocales en abduction.
- 2) La fermeture des lèvres doit être ferme et de durée suffisante pour permettre à la pression de monter dans la cavité orale.
- 3) L'articulation doit rester constante pour éviter une baisse de la PIO par rapport à la PSG.
- 4) L'activité respiratoire doit également rester constante et l'insistance accordée à chaque syllabe doit être égale de manière à obtenir une série de courbes régulières.

Il est impératif d'utiliser une voyelle stable pour permettre l'interprétation des mesures de pression.

4.3.3.3. La pression sous-glottique estimée comme indice de forçage vocal.

Comme nous l'avons vu précédemment dans la littérature, il est fréquent d'associer les mesures de débit d'air à la fuite glottique, les mesures du jitter comme indicateur de raucité ou d'instabilité et de shimmer comme indicateur du souffle (Giovanni et al (1996)^[95]). De la même manière, les mesures de la PSGE pourraient être un indice permettant de rendre compte de la sensation de forçage vocal.

Le forçage vocal est un mode de phonation pathologique caractérisé par une augmentation anormale des tensions à la fois aérodynamiques et musculaires : il correspond à l'augmentation de l'énergie aérodynamique nécessaire pour vaincre la résistance laryngée. La prise en compte de la fuite glottique appréciée par le débit d'air est nécessaire pour estimer la résistance du conduit laryngé.

Cependant, dans le cadre des dysphonies, il peut exister des situations où la fuite est augmentée en raison de la configuration glottique tandis que la PSGE est

augmentée pour compenser cette fuite. De ce fait, la mesure de la PSGE semble utile pour apprécier le forçage vocal Giovanni (1998)^[98].

Plusieurs études ont démontré que la PSGE est plus élevée chez des sujets dysphoniques que chez des sujets témoins :

Zeitels (1997)^[361] a effectué une analyse vocale objective pré et postopératoire chez 20 patients (dont 19 femmes) souffrant d'œdème de Reinke. La stroboscopie préopératoire montrait que la couche superficielle de la lamina propria était dilatée. D'un point de vue acoustique, les sujets présentaient une fréquence fondamentale anormalement basse (123 Hz), et une PSGE anormalement élevée (9.7 cm H₂O²²). La présence d'une PSGE élevée reflète la difficulté pour le sujet de faire vibrer les cordes vocales puisque celles-ci ont une masse plus importante. L'auteur constate également que cette PSGE élevée contribuerait à la dilatation de la couche superficielle de la lamina propria. Après intervention chirurgicale, les patients présentaient une fréquence fondamentale plus élevée et une nette amélioration vocale. Cependant, ils continuaient à produire une PSGE élevée accompagnée d'une tension supra-glottique importante. L'auteur fait alors l'hypothèse que la persistance de ce comportement vocal hyperfonctionnel pourrait contribuer à la réapparition d'une pathologie vocale s'il n'est pas modifié par une rééducation vocale.

Jiang et al (1999)^[154], ont comparé la PSGE de 13 patients consultant pour polype sur les cordes vocales avec 11 sujets témoins. La pression intra-orale et l'intensité ont été enregistrées. Les mesures obtenues pour le groupe contrôle étaient respectivement de 2.38, 2.67 et 2.98 cm H₂O de pression à une intensité de 75, 80 et 85 dB. Les mesures obtenues pour les sujets présentant un polype étaient respectivement de 4.79, 5.85 et 7.37 cm H₂O de pression à une intensité de 75, 80 et 85 dB. Les auteurs concluent en disant que : chez les sujets présentant un polype les valeurs de la pression étaient significativement plus élevées pour chacune des trois tâches (p=0.013 ; p=0.017 ; p= 0.01).

Giovanni 1998^[98] a également enregistré les voix de sujets dysphoniques à l'aide de l'appareillage EVA.

²² **1 hpa** = 1.0195 cm H₂O.

Dans une première recherche, 27 sujets dysphoniques de sexe féminin et 20 sujets féminins ne présentant aucun antécédent de troubles vocaux ont été enregistrés. La tâche des sujets consistait à émettre 8 séries de 10 [pa] à hauteur et à intensité différentes. Il leur était demandé de les prononcer à hauteur usuelle et intensité "normale" ; "faible" (mais non chuchotée), "moyennement faible" ; "forte" et "moyennement forte" ; ensuite à intensité normale mais hauteur usuelle, hauteur de plus de 3 demi-tons ; de plus de 6 demi-tons et de plus de 9 demi-tons.

Les résultats ont démontré que les sujets dysphoniques utilisent une pression sous-glottique plus élevée que les sujets témoins dans toutes les conditions d'intensité et de fréquence.

Tableau 42 : Valeurs moyennes de la PSGE (hpa) mesurées dans différentes conditions d'intensité rapportées par Giovanni (1998)^[98].

	Int.* faible	Int*. moy faible	Int.* normale	Int.* forte moy	Int.* forte
Témoins n=20	3.89 (1.21)**	4.89 (0.96)**	6.10 (1.19)**	7.79 (1.26)**	9.89 (2.39)**
Patients n=27	5.87 (2.14)**	6.81 (2.59)**	8.25 (2.94)**	9.90 (3.15)**	12.31 (3.69)**
P	<0.001	0.003	0.010	0.013	0.039

Légende : *Int. : intensité ; ** : écart type entre parenthèses.

Tableau 43 : Valeurs moyennes de la PSGE (hpa) mesurées en fonction de différentes hauteurs tonales rapportées par Giovanni (1998)^[98].

	Fu	Fu +3	Fu +6	Fu +9
Témoins n=20	6.10 (1.19)*	6.91 (1.14)*	7.26 (1.36)*	8.3 (1.46)*
Patients n=27	8.25 (2.94)*	9.93 (3.32)*	11.38 (4.98)*	11.44 (3.16)*
P	0.016	<0.001	<0.001	0.001

Légende : Fu est le fondamental usuel, +3 = plus 3 demi-tons, +6 = plus 6 demi-tons, +9 = plus 9 demi-tons ; * écart type entre parenthèses.

Dans une deuxième recherche, Giovanni et al (1996)^[96] ont comparé les mesures de la PSGE chez des sujets ayant subi une laryngectomie partielle de Tucker et chez des sujets témoins, dans le cadre d'un protocole multiparamétrique incluant un certain nombre de mesures acoustiques.

Les résultats confirment l'intérêt des mesures aérodynamiques puisque aussi bien le débit d'air buccal que la PSGE étaient significativement plus élevés chez les sujets

opérés (valeur pour le débit d'air buccal chez les sujets sains = 169 cm³/S et chez les sujets opérés = 400 cm³/S ; valeur pour la PSGE chez les sujets sains = 6.2 hpa et chez les sujets opérés = 23 hpa).

Les auteurs concluent que la fermeture glottique des patients opérés étant mauvaise, l'augmentation de la PSGE correspond à l'effort vocal produit pour parvenir au seuil de vibration de la muqueuse malgré la fuite d'air.

La mesure de la PSGE permet de différencier les sujets dysphoniques des sujets témoins et s'intègre progressivement à la pratique clinique.

4.3.3.4. La pression sous-glottique estimée en fonction de l'âge.

- **Chez l'enfant**

La résistance du conduit d'air laryngé pendant la phonation est la mesure de l'opposition effective du larynx contre le flux d'air. Dans la littérature, plusieurs études ont utilisé la méthode clinique de Smitheran et Hixon pour estimer la résistance du conduit d'air laryngé chez l'enfant. Les études (Stathopoulos et al (1985)^[297] , (1993)^[298], Subtelney et al (1966)^[300], Bernthal et al (1978)^[20], Brown (1979)^[30]) ont montré que les enfants présentent une valeur plus élevée de la résistance laryngée que l'adulte, bien que quelques variations soient possibles en fonction de l'âge de l'enfant (Zajac et al (1998)^[360]).

Les avis sont néanmoins partagés.

Netsell et al (1994)^[215] ont montré que la PIO était fonction inverse de l'âge, soit plus l'âge augmente, plus la PIO diminue. Leurs études ont rapporté une pression de 12 cm H₂O chez des enfants âgés de 3 à 5 ans et une pression de 7.7 cm H₂O chez des enfants de 9 à 12 ans. Ces résultats sont corroborés par Stathopoulos^[298]. Il rapporte une pression de 8.7 cm H₂O en voix conversationnelle chez des enfants de 10 ans. Selon Zajac et al (1998)^[360], les enfants ont tendance à utiliser une PIO plus importante que les adultes parce qu'ils parlent à un niveau d'intensité plus élevé. Stathopoulos attribue ce phénomène à une stratégie utilisée par l'enfant pour pallier l'infériorité de ses compétences linguistiques face à l'adulte.

A une puissance conversationnelle normale les valeurs de PSGE des enfants atteignent entre 2 et 4 fois les seuils de valeurs prédites. A une puissance conversationnelle forte, entre 4 et 8 fois les seuils de valeurs prédites. A une puissance conversationnelle normale et des hauteurs tonales habituelles, les PSGE des hommes sont plus basses que celles qu'on a observées chez les enfants, mais similaires à celles des femmes adultes. Le niveau de pression sonore (SPL) à une phonation douce et forte est plus basse que celle que fournit le phonétogramme chez les enfants et les femmes adultes. A intensité et hauteur tonale habituelles, les valeurs de SPL sont similaires chez les enfants et les femmes adultes. En doublant la PSGE, le niveau de pression sonore augmente en moyenne de 10.5 dB. A l'opposé, Arkebauer et al (1983)^[8] rapportent que les enfants ont une pression sous-glottique inférieure à celle des adultes.

Hirano et al (1983)^[123] suggèrent une hypothèse intéressante quant à l'explication d'une pression sous-glottique plus élevée chez les enfants. La morphologie des cordes vocales des enfants diffère de celle des adultes. Dans la période précédant la mue, la structure des tissus connectant la muqueuse glottique à la structure de base est plus lâche. La lamina propria n'est pas totalement différenciée des autres couches. Plus encore, le rapport entre l'épaisseur de la muqueuse et la portion membraneuse des cordes vocales est de $\frac{1}{2}$ chez les nouveaux-nés et de $\frac{1}{10}$ chez les adultes. La longueur des cordes vocales à l'âge de 10 ans est de 13-16 mm, mais de 21 à 29 mm pour l'adulte chez qui la portion membraneuse fait 12 à 15 mm.

Stathopoulos et al (1993)^[298], Arkebauer et al (1983)^[7] et Trullinger et al (1983)^[330] ne trouvent aucune différence selon le sexe pour les enfants de 8 à 10 ans. D'un point de vue théorique, il n'est pas étonnant que les enfants présentent une résistance laryngée plus importante que les adultes. En effet, les enfants possèdent des dimensions laryngées plus petites et peuvent utiliser une PSGE plus élevée (comme le montre la PIO) pour atteindre le niveau d'intensité voulu. Par définition, si la PSGE est maintenue constante, un espace laryngé plus petit va restreindre le flux d'air et augmenter la résistance du conduit d'air laryngé.

Tableau 44 : Valeurs de PSG (cm H₂O) en fonction de différentes tâches rapportées par 3 auteurs.

Auteurs	Années	Tâches (consonnes)	SPL*	PSGE (cmH ₂ O)
Subtelney et al ^[300]	1966	P		9.66 (3.34)**
		B		6.57 (1.97)**
		T		9.35 (2.76)**
		D		6.78 (1.73)**
Bernthal et al ^[20]	1978	P		7.1 (1.7)**
		B		5.4 (1.7)**
	(Adolescents)	P		7.7 (1.5)**
		B		4.7 (1.5)**
Stathopoulos et al ^[298]	1993 (4-8ans)	Conversation	Normal	6.2
	(10 ans)	Conversation	Normal	8.7

Légende : *SPL : "Sound Pressure Level" ou niveau de pression sonore ; ** : déviation standard entre parenthèses.

- **Chez l'adulte.**

Chez l'adulte, il apparaît que la PSGE augmente avec l'intensité et plus légèrement avec la hauteur tonale. Dans un travail de Schutte, la PSGE est mesurée entre 4.3 hPa et 14.5 hPa en fonction de l'intensité. Dans les mêmes conditions, Sawashima et al (1987)^[271] retrouvent des valeurs entre 2 et 15 hPa (Giovanni (1998)^[98]), Södersten et al (1995)^[288].

Les résultats de la littérature relatifs aux différences entre les hommes et les femmes sont discordants.

Pour Holmberg et al (1988)^[137], hommes et femmes ne diffèrent pas significativement. Cependant, Higgins et al (1991)^[120] dans une étude ont comparé les aspects aérodynamiques et temporels de 21 hommes (âge moyen = 24 ans) et 21 femmes (âge moyen = 23 ans). Tous les sujets doivent produire le mot "hamper" à intensité constante. L'analyse de la PIO sur le segment [p] montre que les hommes utilisent un niveau de PIO plus élevé que les femmes. Les résultats obtenus en moyenne sont de 5.9 cm H₂O de pression chez les hommes et de 4.7 cm H₂O de pression chez les femmes. Les auteurs attestent que la présence d'une pression plus élevée chez les hommes est confirmée par d'autres études. Ainsi, Stathopoulos et al (1985)^[297], Andreassen et al^[5], Netsell et al (1994)^[215] ont tous rapporté une PIO plus élevée chez les hommes que chez les femmes.

Cette élévation de pression chez les hommes pourrait être expliquée par la plus grande élasticité de leurs poumons qui permettrait un flux d'air et une pression plus importants chez l'homme.

Tableau 45 : Valeurs de PSGE (cm H₂O) en fonction des consonnes P ,B ,T ,D.

Auteurs	Années	Tâches	PSG (cmH ₂ O)	
			Hommes	Femmes
Subtelney et al ^[300]	1966	P	6.43 (1.07)*	7.52 (2.17)*
		B	4.34 (1.42)*	6.05 (1.82)*
		T	6.18 (1.29)*	7.44 (1.92)*
		D	4.52 (1.60)*	6.67 (1.91)*
Bernthal et al ^[20]	1978	P	6 (1)*	
		B	3.3 (1.4)*	

Légende : * :déviati on standard entre parenthèses.

Tableau 46 : Valeurs de PSG (cm H₂O) en fonction de l'intensité.

Auteurs	Années	SPL *	PSG (cm H ₂ O)
Stathopoulos et al ^[297]	1985	66	6
		70	8.5
		75	10.7
Mc Allister et al ^[193]	1998	59.6	3.6
		78.3	5.9
		81.6	12.5

Légende : SPL* : "Sound Pressure Level" ou niveau de pression sonore.

4.3.3.5. La pression sous-glottique estimée en fonction de l'entraînement.

Des études ont pu mettre en évidence que la PSGE varie de façon très précise pendant le chant. Rubin et al (1967)^[262] ont montré que les chanteurs augmentent leur PSGE non seulement pour accroître leur intensité vocale mais aussi la fréquence fondamentale.

Plusieurs auteurs (Van den Berg et al (1957)^[332], Isshiki (1959)^[145], Ladefoged (1963)^[173], Ohman et al (1968)^[219], Lieberman et al (1969)^[182], Hixon et al (1971)^[129], Baer (1979)^[13], Rothenberg et al (1986)^[259], Baken (1987)^[15]) rapportent que la pression est plus faible dans les fréquences hautes du registre modal, intermédiaire

dans les fréquences basses du registre modal, et plus forte chez les falsettistes. Pour Titze (1989)^[321], l'effet de la pression sur une fréquence fondamentale basse est plus importante, les cordes vocales étant plus relâchées ; avec l'augmentation de la PSGE, elles peuvent se tendre. A haute fréquence, la pression est toujours élevée. Par contre, Isshiki^[145] observe que lors d'un passage de registre, du mode modal au mode falsetto, la pression transglottique peut entraîner une diminution de la fréquence, notamment dans le cas du passage aux hautes fréquences. Atkinson (1978)^[10] note que la pression est le facteur physiologique principal dans le contrôle du fondamental, plus précisément pour des valeurs inférieures à 100 Hz. Pour des valeurs de fondamental supérieures à 120 Hz, les muscles crico-thyroïdiens jouent un rôle primordial. Le modèle mathématique de Farley a montré que l'effet de la pression sous-glottique sur le fondamental varie en fonction de l'action des muscles crico-thyroïdien et thyro-aryténoïdien.

Selon Titze (1992)^[323], la raison pour laquelle la PSGE peut être associée à la fréquence est que la tension des cordes vocales nécessite une pression plus élevée que leur relâchement. La pression pulmonaire doit donc être très précise dans le chant parce que la PSGE peut modifier non seulement l'intensité mais aussi la fréquence fondamentale. Dès lors, si le chanteur ne fournit pas une pression adéquate, la justesse du chant peut être altérée, d'où la nécessité pour le chanteur d'un entraînement rigoureux, chaque changement de fréquence étant accompagné d'un changement de pression précis (Sundberg et al (1993)^[309]).

Sundberg et al (1993)^[310] analysent les productions chantées à hauteurs tonales et intensités différentes de 10 chanteurs d'opéra. Ils observent qu'en moyenne, ils doublent la pression lorsqu'ils augmentent leur fréquence d'une octave. Cette pression doublée entraîne une augmentation du niveau de pression sonore de 8 à 9 dB pour une phonation normale, moins si le mode de phonation est stressé.

Cleveland et al (1997)^[39] déterminent les pressions sous-glottique de 6 chanteurs de variété. Il compare les valeurs de pression sous-glottique et les valeurs de pression sonore des chanteurs de variétés à celles des chanteurs classiques. Les mesures sont obtenues sur [p] durant la parole et le chant. Il observe des pressions sous-glottiques variant de 5 cm H₂O à 55 cm H₂O. Un changement dans la pression

sous-glottique est associé à un changement de puissance. Même si une variation inter-sujets est importante, la pression sous-glottique durant une phonation forte est à peu près de deux fois celle obtenue lors d'une phonation douce. Les courbes de la pression sous-glottique estimée varient systématiquement avec la hauteur tonale. Les chanteurs de "country" augmentent la PSGE en fonction de l'intensité et de la hauteur tonale tout comme les chanteurs d'opéra. Durant la parole et le chant, les chanteurs de variétés semblent employer une PSGE similaire à niveau de pression sonore équivalente mais, leur PSGE est plus élevée que celle des chanteurs d'opéra ; par contre, leur niveau de pression sonore n'est pas aussi élevé que celle des chanteurs d'opéra pour une même PSGE. Cleveland l'explique par le fait que les chanteurs de "country" chantent avec un degré d'adduction plus important que les chanteurs d'opéra. Les chanteurs classiques de part leur formation apprennent à gérer leur voix, ce qui les conduit à chanter sans dommage pour le mécanisme vocal. L'utilisation de la PSGE dans le chant apparaît foncièrement différente de son utilisation en voix parlée. En effet, Ohala (1990)^[218] a montré qu'à une intensité déterminée, la PSGE apparaît constante en voix parlée. De plus, l'intensité et la fréquence apparaissent interdépendantes. Ceci montre bien que les exigences imposées à l'appareil respiratoire sont très différentes en voix parlée ou chantée.

Chapitre 5

Notre Contribution

Une première étude est décrite au chapitre 3. Elle concerne l'analyse perceptuelle du vibrato. Elle complète l'étude 4 en validant les paramètres d'une échelle créée pour l'analyse du vibrato.

Ce chapitre comporte dès lors 6 études concernant l'application des mesures objectives explicitées au chapitre 4 soit l'application de la mesure des paramètres acoustiques et aérodynamiques sur des groupes de sujets sains et pathologiques de toutes tranches d'âge (enfants, adultes, personnes âgées) avec ou sans entraînement vocal spécifique.

1. L'application de l'analyseur spectral à la voix chantée et la comparaison des sujets en fonction de leur entraînement vocal (étude 2).
2. L'établissement de valeurs de référence aérodynamiques et acoustiques chez 40 sujets âgés et la comparaison de leurs résultats à une population de sujets jeunes (étude 3).
3. L'analyse du vibrato sur une population de 30 chanteurs tant sur le plan des mesures objectives fournies par le M.D.V.P. que sur le plan des mesures perceptuelles, à l'aide d'une échelle spécifique (étude 4).

L'établissement de valeurs de référence pour les stations E.V.A. 1 et 2 :

4. L'application des mesures aérodynamiques et acoustiques à l'aide de la station E.V.A. 1 à une population d'enfants âgés de 4 à 9 ans appariés en fonction de la pratique ou non du chant en chœur (étude 5).
5. L'application de ces mêmes mesures à une population de sujets dysphoniques souffrant de paralysie unilatérale laryngée et la corrélation des résultats à ceux de l'échelle perceptuelle G.I.R.B.A.S. (étude 6).
6. L'établissement de valeurs de référence acoustiques et aérodynamiques à l'aide de la station E.V.A. 2 en tenant compte cette fois de la mesure de la pression sous glottique estimée à partir 88 sujets sains qui ont servi de référence à l'étude de 15 femmes présentant une pathologie vocale (étude 7).

5.1. ÉTUDE 2 : APPLICATION DE L'ANALYSE SPECTRALE À LA VOIX CHANTÉE.

5.1.1. Introduction.

C'est l'intérêt tout particulier que nous portons à la voix chantée qui nous a conduit à cette étude. L'examen de la voix implique une mesure quantitative à laquelle répond le quotient phonatoire et le phonétogramme, ainsi qu'une mesure qualitative dont l'objectivation a fait l'objet de nombreux travaux explicités dans la partie théorique. L'analyse spectrale permet d'apprécier de manière objective la qualité vocale. Ce travail poursuit les travaux concernant l'utilité de l'analyseur spectral en phoniatrie. Les recherches précédentes mettaient en évidence l'utilité de ce matériel pour discriminer les voix normales et pathologiques grâce au repérage d'indices pathologiques, et de contrôler l'efficacité d'un traitement vocal ou d'une intervention chirurgicale. En ce qui nous concerne, notre contribution porte sur la voix chantée non pathologique chez des sujets dont la profession exige un rendement vocal maximal, et un état laryngé parfait.

Cette étude englobe deux volets le formant du chanteur tel que décrit par Sundberg (1970)^[303], (1977)^[305], (1991)^[308] et l'observation du vibrato.

5.1.2. Matériel et méthode.

L'appareil utilisé est l'analyseur fréquentiel 3561 A de Hewlett-Packard^[117]. Nous avons opté pour un examen court et rapide (environ 15 minutes). La population est constituée de 60 sujets appariés en fonction de leur niveau d'étude de chant classique. Vingt sujets forment le groupe contrôle et ne pratiquent pas l'étude du chant. Vingt autres poursuivent des études de chant en académie. Les vingt derniers ont terminé le Conservatoire et se destinent à la carrière de chanteur. Les sujets émettent successivement 3 voyelles [a¹], [u], [i] et un [a²] à 440 Hz à 30 cm du micro (modèle Beyer m260 200 Ohms), dans une cabine anéchoïde. La pertinence de ces

trois voyelles a fait l'objet de recherches (Sundberg (1970)^[303]). Durant l'expérience, les soprani éprouvent des difficultés à émettre un son [a¹] chanté sur leur fréquence fondamentale. Elles nous disent que cette émission ne correspond pas à leur tessiture. Aussi, pour éviter cet inconvénient, nous avons ajouté l'émission d'un [a] à 440 Hz. L'échelle de fréquence est déterminée en fonction du formant du chanteur soit 0 à 4000 Hz, l'intensité est exprimée en décibel volt.

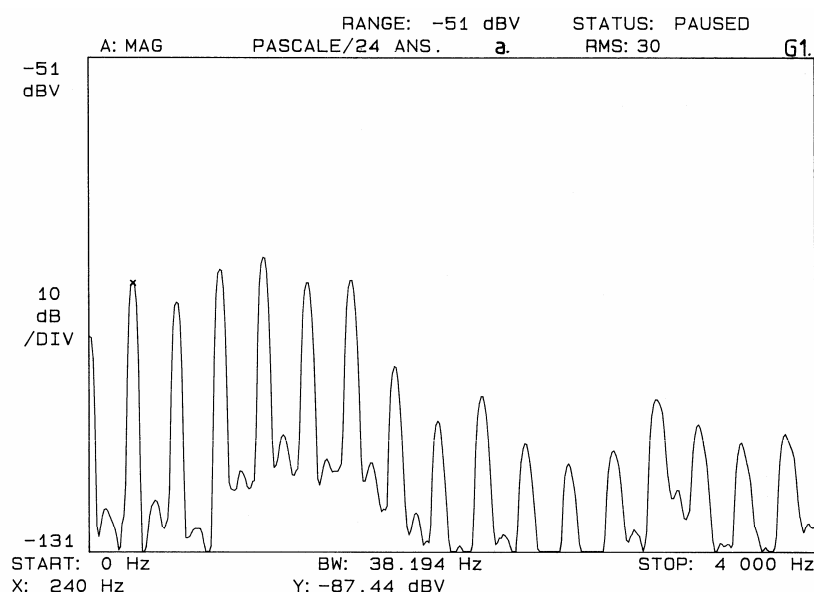


Figure 27 A : Représentation spectral du son [a] émis par une femme non-chanteuse. La fréquence fondamentale est de 240 HZ, l'intensité de -87.44 dBV.

5.1.3. Résultats.

Le tableau 47 montre l'identification du formant du chanteur dans les trois groupes. Le formant du chanteur et le vibrato sont mieux identifiés dans le groupe 1 sur [a¹] et [i], pour le groupe 2 sur [a¹], [a²], et [i]. Le vibrato est identifié dans le groupe 3 sur toutes les voyelles, le formant du chanteur, lui, apparaît mieux sur le [a²].

Tableau 47 : Observation du formant du chanteur dans les 3 groupes pour chaque voyelle émise.

Groupes	[a ¹]	[u]	[i]	[a ²]	Totaux
1	4	1	4	---	9/60
2	7	2	8	8	25/80
3	18	11	11	20	60/80

Le tableau 48 résume la présence du vibrato. Dans le groupe 1. cinq sujets ont un vibrato qui est plus marqué sur [a²] et [i]. Il apparaît sur les derniers harmoniques du tracé. Dans le groupe 2 et ce, plus fréquemment sur le [a²] et le [i], nous enregistrons le vibrato chez 13 sujets. Il apparaît autour des 3000 Hz. Dans le groupe 3. presque tous les chanteurs présentent un vibrato ; pour certains d'entre eux, il apparaît déjà à 800 Hz.

Tableau 48 : Observation du vibrato dans les 3 groupes pour chaque voyelle émise.

Groupes	[a ¹]	[u]	[i]	[a ²]	Totaux
1	3	2	4	---	9/60
2	9	4	8	5	26/80
3	14	13	14	14	55/80

Tableau 49 : Observation du formant et du vibrato en fonction de la tâche et du groupe (Test binomial).

Formant	[a ²] vs [u]	[u] vs [i]	[a ²] vs [i]
Groupe 1	NS	NS	NS
Groupe 2	NS	NS (0.070)	NS
Groupe 3	0.016	NS	NS (0.070)
Vibrato			
Groupe 1	NS	NS	NS
Groupe 2	NS	NS	NS
Groupe 3	NS	NS	NS

La différence d'acquisition du formant est fonction du groupe de façon significative. Plus le sujet pratique le chant, plus nous observons la présence d'un formant.

Tableau 50 : Comparaison de l'acquisition du formant du chanteur et du vibrato en fonction des 3 groupes.

	Test de Cochran Gradation 1<2<3<	Test exact de Fischer		
Formant		1 vs 2	2 vs 3	1 vs 3
[a ²]	<0.001	NS	<0.001	<0.001
[u]	<0.001	NS	0.006	0.001
[i]	<0.023	NS	NS	0.048
Vibrato				
[a ²]	<0.001	NS (0.082)	NS	0.001
[u]	<0.001	NS	0.010	<0.001
[i]	<0.001	NS	NS	0.004

5.1.4. Discussion et Conclusion.

Le formant du chanteur est observé plus fréquemment chez les chanteurs que chez les non-chanteurs. Selon Sundberg (1987)^[306], le formant résulte de l'aperture buccale, souvent pratiquée par les chanteurs, notamment pour atteindre les fréquences aiguës. Une telle ouverture, exagérée par rapport à la voix parlée, facilite l'abaissement du larynx. Le pharynx ainsi largement ouvert suit mieux les sons glottiques pour prendre place dans les résonateurs pharyngés et buccaux, provoquant ainsi des sensations vibratoires au niveau des os du crâne, excellent point de référence pour une bonne diffusion vocale. Comme Sundberg (1987)^[306], nous trouvons que chez les sujets appartenant aux groupes 1 et 2 [figures 27 B et 27 C]. le formant est plus présent chez les hommes, principalement sur des fondamentaux bas [Figures 27 D et 27 E].

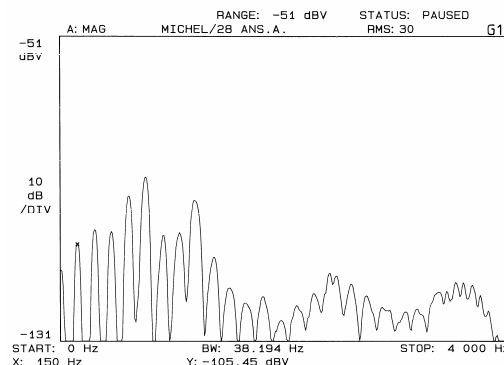


Figure 27 B : Représentation spectrale du son [a] émis par un homme non-chanteur. La fréquence fondamentale est de 150Hz et l'intensité de -150.45 dBV.

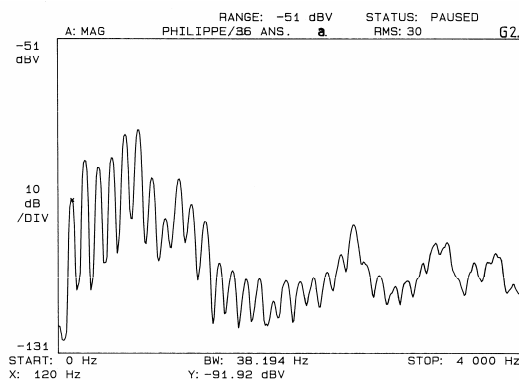


Figure 27 C : Représentation spectrale du son [a] émis par un chanteur en académie. La fréquence fondamentale est de 120 Hz et l'intensité de -91.92 dBV. Nous notons une amplification de la courbe dont le centre du formant se situe à 2562 Hz. Un vibrato naissant se distingue dans les derniers harmoniques.

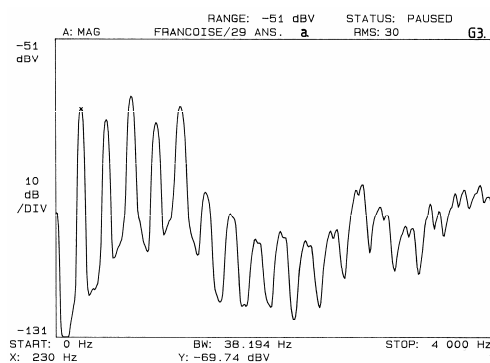


Figure 27 D : Représentation spectrale du son [a] émis par une mezzo. La fréquence fondamentale est de 230 Hz et l'intensité de -69.74 dBV. Nous notons une amplification de la courbe dont le centre du formant est situé à 2818 Hz, ainsi qu'un vibrato apparaissant au 7^{ème} harmonique.

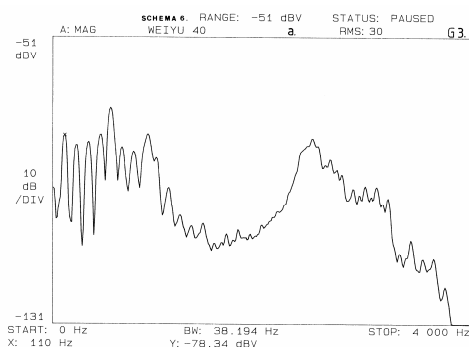


Figure 27 E : Représentation spectrale du son [a] émis par une basse. La fréquence fondamentale est de 110 Hz et l'intensité de -78.34 dBV. Nous notons une amplification très claire de la courbe soit un formant du chanteur dont le centre est situé à 2469 Hz. Le vibrato est présent à partir du 8^{ème} harmonique.

Nous observons un formant du chanteur chez les femmes essentiellement dans les groupes 2 et 3 [Figure 27 F], tant chez les soprani que chez les alti.

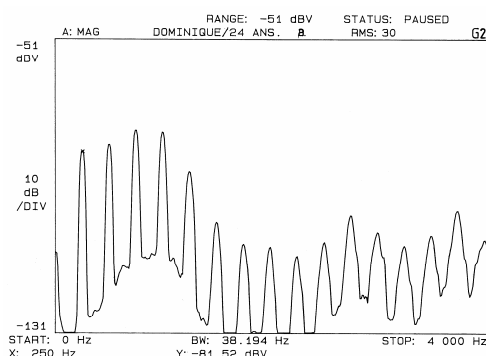


Figure 27 F : Représentation spectrale du son [a] émis par une mezzo étudiant en académie. Le fondamentale est de 250 Hz et l'intensité de -81.52 dBV.

Le fait que le formant est moins apparent chez les chanteuses peut être expliqué comme suit :

Les femmes chantent une octave plus haut que les hommes. Il semblerait dès lors inutile qu'elle développe un formant. Leur tessiture plus aiguë leur permet d'être facilement au-dessus de l'orchestre particulièrement pour les soprani. Ce n'est pas le cas des chanteurs qui le développeraient pour parer à un accompagnement bruyant. Le formant du chanteur serait dès lors le résultat de la nécessité d'être entendu, mais également de son travail et du talent qu'il déploie.

L'emploi du son [a] est utile en analyse spectrale. Le fondamental du son [a] s'avère être plus bas, il recouvre un maximum d'harmoniques pour une même échelle de fréquences. Nos résultats montrent que le vibrato et le formant apparaissent en plus grand nombre sur [a]. Son spectre montre autant d'informations dans les fréquences graves que dans les fréquences aiguës. Le son [u], par contre, ne facilite pas l'apparition de ces critères. Seuls les chanteurs chevronnés seraient capables de faire émerger le formant. Nous le considérons donc comme étant intéressant dans la mesure ou sur une étendue plus fine, il nous permet d'évaluer les capacités d'un chanteur, et de son placement vocal. Le son [i] n'apporte pas beaucoup plus d'éléments que le [a], si ce n'est une meilleure observation des harmoniques aigus.

Cependant, il est assez étonnant de constater que ce son facilite peu l'apparition du formant chez les chanteuses du groupe 3. alors qu'il est souvent choisi pour sa facilité d'émission : il transmet aisément des vibrations au niveau des os du crâne. Il ne nécessite pas une grande intensité pour apparaître. Ceci pourrait s'expliquer par leur tessiture plus aiguë, puisque les seules présentant un formant sur le [i] sont les mezzo-soprani.

Pour identifier le vibrato, nous avons mis en relation ce que nous percevions auditivement et l'image spectrale correspondante. Les vibrati relevés sont essentiellement ceux qui se marquent sur les courbes spectrales par un élargissement des pics et une configuration en dents de scie. Les résultats semblent logiques. Plus nous avons à faire à des chanteurs professionnels, plus le vibrato s'observe sur les premiers harmoniques, s'affirme et se régularise. Le fait que nous puissions aisément l'identifier chez les chanteurs du groupe 3 et qu'il s'identifie sur toutes les voyelles traduit une aisance dans la production vocale. Il est moins présent dans le groupe 2 vraisemblablement parce que les chanteurs sont en cours d'apprentissage et que le vibrato (quoique présent de manière naturelle chez certains apprentis chanteurs) n'a pas encore eu le temps de se développer de façon régulière et de s'amplifier. Une autre explication serait que la moyenne choisie (30 RMS²³) n'a pas permis de le faire apparaître de manière évidente sur la courbe. Cela traduirait cependant que le vibrato est moins régulier et moins stable que chez les professionnels. Les quelques vibrati observés dans le groupe 1 se marquent sur les derniers harmoniques. Leur présence signifierait qu'ils sont naturels et inhérents à la voix.

²³ *Le **RMS**: correspond à "root-mean-square" ou encore niveau efficace mesuré par la racine de la moyenne des carrés des amplitudes instantanées. Plus l'étendue fréquentielle est grande, plus cette fonction décroît rapidement à l'écran. En la fixant à 30, la durée d'extraction du son est de plus ou moins 10 secondes. Ce temps phonatoire peut paraître court, mais certains sujets du premier groupe n'auraient pas pu sonoriser au-delà de 10 secondes.*

5.2. ÉTUDE 3 : VALEURS AÉRODYNAMIQUES ET ACOUSTIQUES DE RÉFÉRENCE POUR UNE POPULATION DE SUJETS ÂGÉS DE 60 À 90 ANS.

5.2.1. Introduction.

Grâce aux progrès scientifiques dans la prévention et l'amélioration des conditions de vie, l'espérance de vie est passée de 31 ans au début du 19^{ème} siècle à 76 ans actuellement (Swine (1992)^[314]). Prévoir les troubles vocaux dus au vieillissement est tout aussi important que de prévenir certaines maladies chroniques comme les maladies cardio-vasculaires, les affections respiratoires, ou encore les affections du système ostéoarticulaire, du système digestif, etc. Ces dépistages précoces permettent aux personnes âgées "d'entretenir" leur outil vocal afin de sauvegarder une communication optimale avec leur entourage.

A l'audition et de manière subjective, la voix d'une personne âgée est sensiblement différente de celle d'une personne jeune, notre étude a pour but de mesurer le rendement et la qualité vocales chez la personne âgée de sexes masculin et féminin afin d'apprécier en quoi diffère les voix âgées des voix plus jeunes et d'objectiver les changements vocaux qui les caractérisent.

5.2.2. Matériel et Méthode.

L'étude a porté sur une population de 40 personnes - 30 femmes et 10 hommes de 70 à 90 ans et n'ayant pas de pathologie autres que celles liées à leur âge. Nous avons demandé au sujet d'émettre à 5 reprises un [a] long et nous avons retenu le meilleur résultat. L'intensité (60 dB) était contrôlée par un décibel-mètre placé près du microphone. Le microphone (AKG C-Type 1000 S, Vienna, Austria), connecté à un enregistreur (D.A.T., Germany ; TEAC, DA-P20. Musashino-Shi, Japan) était placé à 30 cm de la bouche du sujet. Le signal enregistré a ensuite fait l'objet d'une analyse fréquentielle à haute résolution par transformée de Fourier (Hewlett Packard 3561A,

Palo Alto, Californie^[117]). Les mesures obtenues ont été étudiées sous la forme de deux graphiques portant en ordonnée l'intensité exprimée en dB volt et en abscisse une échelle de fréquence de 0 à 1000 Hz pour le premier graphique et 0 à 5000 Hz pour le second ; la fourchette des fréquences a été choisie pour déterminer le fondamental et les 4-5 premiers harmoniques. Une évaluation qualitative de ces tracés s'est faite au moyen d'une échelle à cinq niveaux.

0- le tracé plat, absence d'harmoniques et de sonorisation.

1- présence de bruit et absence de son.

2- présence de la fréquence fondamentale, pauvreté en harmoniques et présence de bruit.

3- le fondamental est présent ainsi que les harmoniques. Il y a persistance de bruit.

4- présence du fondamental et des harmoniques avec absence de bruit.

[cfr figure 20].

Cette classification qualitative résulte d'études antérieures ^[249. 252. 253]. Les résultats sont moyennés selon la racine carrée de la moyenne des carrés RMS (root mean square). Pour chaque prise de mesure, 30 échantillons sont analysés. L'analyseur évalue l'amplitude de chaque fréquence (multiple entier de 2.5 Hz pour la prise de mesure sur 1000 Hz et de 10.5 Hz pour la prise de mesure sur 5000 Hz). La moyenne quadratique est calculée par $(\sum A_i^2/30)^{1/2}$.

Grâce à la table traçante de type HP 7475A (Palo Alto, Californie), raccordée à l'analyseur, on peut obtenir les tracés format A4 du phonème prononcé par le sujet. Ce tracé permet d'apprécier la qualité de son vibreur vocal.

Les mesures objectives ont été établies conjointement à l'analyse fréquentielle. Nous avons retenu les trois tests qui nous ont donné des résultats significatifs dans notre étude. Ce sont : la capacité vitale, le temps maximum phonatoire et le quotient phonatoire.

La capacité vocale est mesurée par le spiromètre. Le temps maximum phonatoire est mesuré avec un chronomètre après avoir demandé au sujet d'inspirer profondément et de tenir une voyelle le plus longtemps possible. L'intensité de la fréquence fondamentale parlée est contrôlée à 60 dB. De ces valeurs mesurées 3 fois, la meilleure est retenue. Le temps maximum phonatoire résulte de la capacité vitale divisée par le temps maximum phonatoire mesuré sur [a] (CV/TMP = QP en ml/sec)

(valeurs normales = 150-200 ml/sec). Les valeurs numériques sont résumées par la moyenne des déviations standard, et par les moyennes de l'étendue inter quartile. CV, TMP, et QP sont comparées entre les deux sexes par les Wilcoxon rank sum tests. Nous employons le t de student pour comparer nos résultats avec des paramètres similaires de sujets jeunes, pour lesquels des estimations communes sont extraites d'Inglis et al (1985)^[144] et de Taylor et al (1985)^[316] rapportées au tableau 51.

Tableau 51 : Mesures extraites de l'étude d'Inglis et Taylor sur une population jeune ^[144. 316].

Mesures	Population totale	F	H
CV (ml)	4196 (+/-921)*	3479.2 (+/-405.9)*	4930.0 (+/-666.5)*
TMP (sec)	23.8 (+/-6.6)*	22.8 (+/-4.1)*	24.8 (+/- 8.4)*
QP (ml/sec)	187.6 (+/-61.6)*	157.2 (+/-35.6)*	217.8 (+/-67.7)*

Légende : CV : capacité vitale ; TMP : temps maximum phonatoire ; QP : quotient phonatoire

* déviation standard entre parenthèses.

5.2.3. Résultats.

Sur le plan de l'analyse de fréquence vocale que nous avons défini préalablement, les résultats de notre recherche corroborent ceux rencontrés dans la littérature ^[135].

Tableau 52 : Valeurs moyennes de la capacité vitale, du temps maximum phonatoire, du quotient phonatoire et de la fréquence fondamentale pour les deux sexes.

F (n=30)	Ages	CV (ml)	TMP (sec)	QP	Fo
Range	70-89	100-2900	4-20	17-450	152.5-350
Moyenne	83 (+/-5)*	1600 (+/-600)*	12 (+/-5)*	154(+/-87)*	205 (+/-43)*
Médiane	83 [8]	1600 [500]	12 [9]	152 [108]	197.5 [37.5]
H	Age	CV (ml)	TMP (sec)	QP	Fo
Range	73-89	1400-4400	11-39	74-250	102.5-250
Moyenne	82 (+/-4)*	2700 (+/-700)*	20 (+/-9)*	153 (+/-48)*	158 (+/-44)*
Med	82 [4]	2600 [600]	16[14]	153 [64]	152.5 [37.5]

Légende : F : femmes, H : hommes ; * :déviation standard entre parenthèses ; Med : médiane.

La fréquence fondamentale moyenne chez les femmes âgées diminue alors que chez les hommes âgés, elle augmente. Il s'ensuit que l'écart entre le fondamental moyen des hommes et des femmes âgés est réduit ; le fondamental de l'homme se rapprochant de celui de la femme (fondamental moyen de l'homme âgé est de 158 Hz pour une valeur normale de 120 Hz et le fondamental moyen de la femme âgée est de 205 Hz pour une valeur normale de 225 Hz). La distinction entre la voix masculine et la voix féminine chez les personnes âgées est dès lors moins évidente.

Le pourcentage établi chez les personnes âgées pour chaque niveau de l'échelle montre que les hommes présentent une meilleure qualité vocale que les femmes âgées. En effet, 70% d'entre eux présentent un graphique de niveau 3 alors que 47% des femmes seulement se situent dès ce niveau et 10% des hommes obtiennent un tracé de voix normal alors qu'aucune femme n'atteint ce niveau [Figure 28].

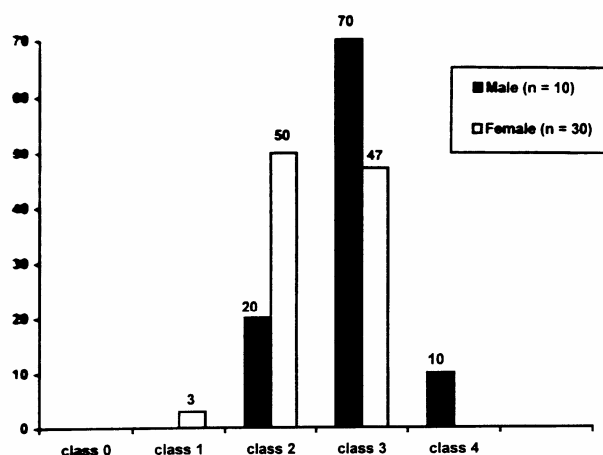


Figure28 :Distribution en pourcentage de la qualité vocale (obtenue à l'analyse spectrale et mesurée sur une échelle à 4 niveaux) des sujets âgés hommes et femmes.

Nous avons comparé les mesures quantitatives entre les hommes et les femmes âgées. D'après le test de Wilcoxon²⁴, nous constatons que la capacité vitale, la temps maximum phonatoire et le quotient phonatoire diffèrent significativement

²⁴ Le **test de Wilcoxon** est un test non paramétrique qui par définition, n'utilise pas les valeurs observées mais les rangs de ces valeurs. Il est indépendant de la distribution des valeurs.

entre les femmes et les hommes (médiane CV=2610 pour les hommes âgés et 1610 pour les femmes âgées ($p<0.001$) ; médiane TMP = 16 sec pour les hommes âgés et 12 sec pour les femmes âgées ($p=0.025$). Le vieillissement influence la capacité vitale et par conséquent la tenue phonatoire. Celles-ci sont significativement meilleures chez l'homme que chez la femme. Nous observons que le quotient phonatoire ne diffère pas significativement entre les sexes (médiane QP= 153 ml/sec chez les sujets âgés et 152 ml/sec chez les femmes âgées). La comparaison de nos résultats chez les sujets âgés avec ceux des sujets jeunes montre une diminution des valeurs sur tous les paramètres pour les deux sexes. Une seule exception chez les femmes pour le quotient phonatoire : la capacité vitale (CV) moyenne des femmes âgées est égale à 46% de celle des femmes jeunes ($p<0.001$) et la CV chez les hommes est égale à 55% de celle des hommes jeunes ; le temps maximum phonatoire (TMP) chez les femmes âgées est égal à 59% de celui des femmes jeunes ($p<0.001$) et la moyenne du TMP chez les hommes âgés est égale à 59% de celle des sujets jeunes ($p<0.001$) ; la moyenne du quotient phonatoire (QP) chez les femmes âgées est égale à 98% de celui des femmes jeunes (NS) et la moyenne du QP chez les hommes âgés est égale à 75% de celui des hommes jeunes ($p=0.018$) [Tableau 53]

Tableau 53 : Comparaison des mesures quantitatives entre les sujets jeunes et âgés.

Paramètres	Sexes	Comparaisons	% jeunes/âgés valeurs moyennes
CV (ml)	F	$P<0.001$	46%
TMP (sec)	F	$P<0.001$	59%
QP (ml/sec)	F	NS	98%
CV (ml)	M	$P<0.001$	55%
TMP (sec)	M	$P<0.001$	59%
QP (ml/sec)	M	$P=0.018$	75%

Légende : CV : capacité vitale ; TMP : temps maximum phonatoire ; QP : quotient phonatoire.

Nous avons comparé les résultats de ces mesures quantitatives chez les sujets âgés avec les sujets jeunes. Nous avons tiré nos valeurs pour les sujets jeunes des tableaux d'Inglis et al (1985)^[144] et de Taylor et al (1985)^[317] et en avons dérivé des estimations communes après avoir vérifié l'homogénéité des deux séries. Le test t de

Student²⁵ a été utilisé dans le but d'observer si les différences sont significatives ou non. On observe un effondrement de 40% et plus de la capacité vitale et de la tenue phonatoire chez les personnes âgées par rapport aux sujets jeunes.

Les femmes âgées ont une capacité physique qui s'amenuise mais leur quotient phonatoire (QP) reste constant par rapport aux femmes jeunes, même si le résultat obtenu n'est qu'indicatif. Par contre, les hommes perdent un quart de leur rendement vocal par rapport aux sujets jeunes.

5.2.4. Discussion et Conclusion.

L'analyse de la fréquence fondamentale (Fo) montre que celle-ci décroît chez les femmes âgées et s'accroît chez les hommes âgés. Cette observation rejoint les résultats rencontrés dans la littérature^[22. 28. 57]. Elle peut s'expliquer par l'apparition d'œdème sous-muqueux au niveau des cordes vocales avec la ménopause chez les femmes et par l'atrophie du muscle vocal chez les hommes^[22. 144].

La difficulté de distinguer à l'audition les voix âgées féminines des masculines est due au fait que le fondamental de l'homme rejoint celui de la femme.

La différence dans la hauteur tonale vocale entre les hommes et les femmes est réduite en relation avec la modification de Fo dû à l'altération physique de la source vocale avec l'âge.

Le pourcentage établi chez les personnes âgées pour chaque niveau de l'échelle montre que les hommes ont une qualité vocale supérieure à celles des femmes du même âge. En effet, 70 % d'entre eux présentent un graphique de niveau 3 alors que 47 % des femmes seulement se situent à ce niveau et 10 % des hommes obtiennent un tracé de voix normal alors qu'aucune femme n'atteint pas ce niveau. Nous avons comparé les mesures quantitatives des hommes âgés avec celles des femmes âgées. On constate que le vieillissement influence la capacité vitale à l'avantage des hommes. Cette différence s'observe déjà chez les sujets jeunes, la capacité vocale d'un homme jeune étant nettement supérieure à celle d'une femme jeune. Inglis et al^[144] mesuraient 3479 ml, et 4930 ml ; et Taylor et al^[317] observaient

²⁵ Le **test de Student** rend compte de la pertinence d'une comparaison établie pour deux moyennes d'échantillons indépendants. Ce test suppose une distribution normale des échantillons et une égalité de leurs variances,

respectivement 3376.2 ml et 4973 ml. On peut donc dire que malgré le vieillissement, la différence entre les sexes se maintient.

Les résultats du temps maximum phonatoire sont significativement meilleurs chez les hommes âgés comparés à ceux des femmes âgées. Ils sont la conséquence d'une plus grande capacité vitale chez les hommes âgés (médiane TMP=16 sec chez les hommes ; 12 sec chez les femmes). D'autre part, le quotient phonatoire ne révèle pas de différence significative selon le sexe (QP pour les hommes (médiane) = 154. pour les femmes = 153).

Nous observons un effondrement de 40 %, voire plus, de la capacité vitale et du temps maximum phonatoire chez les personnes âgées par rapport aux sujets jeunes.

Cela peut s'expliquer par l'atrophie musculaire inhérente à l'âge.

Les femmes ont une capacité physique qui s'amenuise avec l'âge mais le rendement vocal reste bon par rapport aux jeunes femmes, même si ce résultat n'est pas significatif (QP moyen chez les femmes jeunes est de 157ml/sec, chez les femmes âgées de 154 ml/sec). Par contre les hommes âgés perdent 25% de leur rendement vocal par rapport aux sujets jeunes, et ceci de manière significative.

Entre les hommes jeunes et les hommes âgés, le débit expiratoire pour la tenue du [a] est constant avec l'âge alors que chez l'homme il y a proportionnellement une plus grande diminution de la capacité vitale par rapport à la tenue du a (la capacité vitale moyenne pour les hommes jeunes est de 4.9 litres tandis qu'elle est de 2.7 litres pour les hommes âgés ; la tenue maximale du [a] est de 24.8 sec. pour les hommes jeunes et de 19.5 sec. pour les hommes âgés). Aussi, une diminution du quotient phonatoire traduisant une dégradation du rendement vocal en découle.

Il convient cependant d'être prudent dans les conclusions relatives au quotient phonatoire. Il est tout à fait possible en effet que Inglis et al (1985)^[144], et Taylor et al (1985)^[317] dans leurs études aient utilisé des appareils différents des nôtres. Des changements dans le calibrage ou dans la réalisation du test pourraient modifier sensiblement les résultats, notamment le temps pris au chronomètre qui peut varier en fonction de la vitesse de déclenchement de l'expérimentateur.

Néanmoins, les mesures obtenues dans notre étude sont issues de la pratique journalière et sont facilement réalisées. Les appareils requis (spiromètre, chronomètre) ont dès lors une fiabilité à peine influencée par les critères de calibrage.

contrairement au test de Wilcoxon qui lui ne tient pas compte de la forme de la distribution.

En ce qui concerne notre classification de l'analyse fréquentielle, déjà employée dans des études précédentes, cette méthode nous donne maintenant un bon équilibre entre la fiabilité souhaitée et la simplicité nécessaire dans la pratique clinique journalière. Cette classification est employée pour toutes les évaluations vocales (par exemple : la phono-chirurgie, les laryngectomies partielles ou totales, durant la rééducation logopédique). Des cas peuvent être inclus, de l'absence de son (classe 0) à une voix normale (classe 4). Cela pourra être considéré comme démodé quand et si une standardisation plus satisfaisante de ces types d'analyse sera proposée.

Les résultats de notre travail corroborent ceux que fournit la littérature. La fréquence fondamentale moyenne chez les femmes âgées diminue tandis que celle des hommes augmente. Aussi, la distinction entre la voix masculine et la voix féminine chez la personne âgée est dès lors moins évidente. Enfin, les hommes âgés ont un moins bon rendement vocal mais une meilleure qualité du vibreur vocal que les femmes âgées. Le vieillissement influence la capacité vitale à l'avantage des hommes. Cette différence également observée entre les sujets jeunes des deux sexes, se maintient avec l'âge. Le temps phonatoire est également significativement supérieur chez les hommes. Nous constatons une nette diminution des performances des sujets âgés pour la capacité vitale et la tenue du [a] par rapport aux sujets jeunes. Bien que la capacité physique des femmes âgées soit réduite, leur rendement vocal reste fort bon par rapport aux femmes jeunes. Les hommes quant à eux montrent une dégradation du rendement vocal de 25%.

5.3. ÉTUDE 4 : APPLICATION DU M.D.V.P. A LA MESURE DU VIBRATO ET CORRELATION AVEC UNE ECHELLE PERCEPTUELLE.

5.3.1. Introduction.

Depuis quelques années, la consultation en phoniatry recrute de plus en plus d'artistes vocaux, comédiens, chanteurs d'opéra, chanteurs de variété, ventriloques, imitateurs, orateurs professionnels. Cette étude se focalise sur les chanteurs d'opéra. L'étude 2 a permis d'objectiver deux caractéristiques de la voix à savoir le vibrato et le formant du chanteur. Cette quatrième étude se centre uniquement sur le vibrato défini comme une modulation de la hauteur du son et habituellement de l'intensité et du timbre. Il est dépendant de l'état musculaire.

Depuis les années 90. on assiste à un consensus plus élargi sur le plan de la terminologie et de la mesure du vibrato de même que sur les données chiffrées des différents paramètres du vibrato, du trémolo, du trille et du chevrottement. Le tableau 54 répertorie des valeurs de fréquences, d'amplitudes et d'étendues selon différents auteurs. La figure 29 illustre d'un point de vue physico-acoustique les notions de cadences et d'étendues moyennes.

Tableau 54 : Valeurs de cadence et d'étendue du vibrato selon différents auteurs.

Auteurs	Années	Fréquences	Amplitudes	Étendues en ½ Tons
Miller ^[200]	1986	8 cycles/sec		
Rothman et al ^[260]	1987	5 à 6.7 cycles/sec	5 à 7 cycles/s	0.5 à 1.8
Hakes et al ^[107]	1990	5 à 7 cycles/sec		
Rothman et al ^[261]	1990	10 à 13 cycles/sec	3.7/sec	
Sapir et al ^[266]	1993	4 à 5 cycles/sec	2 à 8 cycles/s	¼ à ½
Cleveland ^[38]	1994	5.5 à 7 cycles/sec		1 à 2
Titze ^[327]	1994	5 cycles/sec		
Dejonckere ^[in 63]	1995	4.1 à 6.6 cycles/sec		0.38 à 3.26

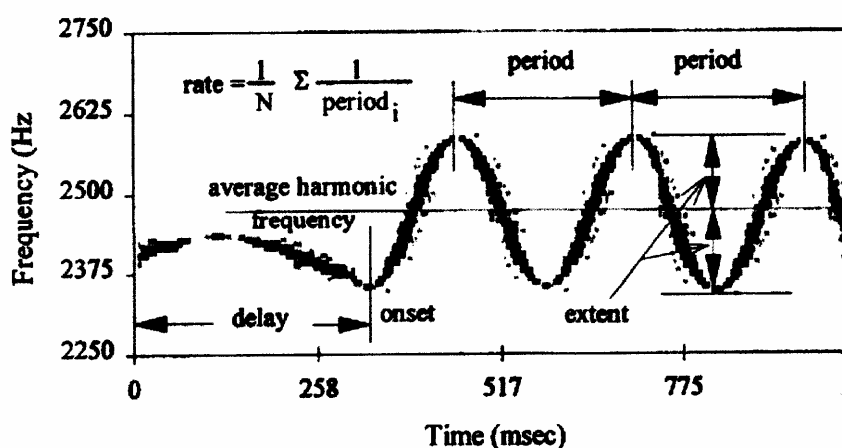


Figure 29 : Mesures de la cadence et de l'étendue moyennes du vibrato, du délai de mise en vibration (voice onset) du vibrato. Mesures obtenues par spectrogramme à bande étroite à l'aide d'une procédure semi-automatique (Sataloff).

Nous avons étudié le vibrato de deux façons, d'une part en analysant la constance de l'oreille humaine dans le jugement d'une trémulation de la voix chantée. Nous avons voulu savoir si un sujet peut ou non être concordant d'une écoute à l'autre et si plusieurs sujets peuvent être concordants entre eux par rapport à un même morceau chanté. D'autre part en sélectionnant des paramètres subjectifs (issus de l'analyse de la constance de l'oreille et mis en évidence par des calculs de concordance et de corrélation) et en les comparant à des paramètres objectifs sélectionnés sur base de la littérature (Miller (1986)^[200]; Rothman et al (1987)^[260]; Rothman et al (1990)^[261]). Les questions qui ont guidé notre étude sont les suivantes :

- L'oreille est-elle constante dans la qualification d'un vibrato? (Les membres du jury répondent-ils de la même façon à la première puis à la deuxième audition? (fidélité Test/Re-test))
- Quelle est la tâche qui obtient la meilleure concordance, tous membres du jury confondus?
- Peut-on faire un lien entre certaines données objectivées par l'analyse au M.D.V.P. et certains jugements qualitatifs du jury d'écoute?

5.3.2. Matériel et Méthode.

L'étude a porté sur 30 chanteurs. Dix sont professionnels, dix semi-professionnels, et dix étudient le chant. Pour chaque groupe, nous avons choisi cinq hommes et cinq femmes entre vingt et cinquante-cinq ans. Aucun de ces sujets ne présentait de pathologie laryngée connue à ce jour. Nous avons demandé à chaque chanteur de réaliser quatre tâches. La première consistait à émettre un [a] tenu vibré et la deuxième un [a] tenu non-vibré. La troisième tâche consistait pour le chanteur à exécuter un extrait de son choix de 18 mesures. Il s'agissait d'un extrait que le chanteur appréciait et maîtrisait. Dans la quatrième tâche, il s'agissait pour chaque chanteur d'apprendre en 15 jours un morceau original composé pour l'occasion par une professionnelle et de le chanter [cfr figure 19]. Ce morceau a été transposé dans toutes les tessitures, il pouvait être chanté par tous les types de voix. Ces troisième et quatrième tâches, nous ont donc permis d'étudier le vibrato dans un contexte musical. Nous avons enregistré 120 échantillons au total.

Pour chacune des quatre tâches, le sujet a eu droit à trois essais, et choisissait celui qui le satisfaisait le plus. Tous les morceaux ont été chantés à capella. Tous chanteurs ont reçu la même consigne avant leur prestation. Nous avons tenu à les rassurer, afin de leur permettre d'être le plus à l'aise possible, et d'éviter ainsi des interférences. Ils savaient qu'ils participaient à une étude dans le cadre d'un mémoire. Dans cette mise en condition préalable, nous les avons assurés que nous ne portions pas un jugement qualitatif et que le seul but était d'améliorer les rééducations vocales. Nous avons choisi un jury d'écoute composé de trois membres. Chacun d'eux était professionnel depuis plusieurs années et prestait un minimum de cinq contrats²⁶ par an. Ils disposaient également d'une expérience pédagogique et d'une bonne connaissance théorique du vibrato. Nous avons fait entendre deux fois les 120 échantillons présentés en ordre aléatoire à chaque membre du jury. Nous les avons informés qu'ils entendraient 240 échantillons vocaux et que leur tâche consistait à porter un jugement sur la trémulation et l'oscillation. Nous n'avons donc pas précisé qu'il s'agissait d'entendre deux fois la même série d'échantillons. Afin d'éviter les interférences, nous avons veillé à ce que l'expérimentateur qui travaillait avec le jury

²⁶ Nous entendons par contrat, des engagements rémunérés pour des prestations grand public.

d'écoute soit différente de celle qui avait enregistré les chanteurs. Ainsi, aucun commentaire, ne pouvait être fait durant l'écoute. Il était indispensable de nous accorder sur la terminologie. C'est pourquoi, nous avons décidé d'une phase préliminaire où nous informions le jury de ce que nous entendions par vibrato, trémolo, chevrottement et son blanc. Ces informations étaient appuyées par des instructions écrites [annexe 9]. Nous avons rencontré chaque membre du jury et au minimum trois fois, afin d'éviter toute saturation auditive. Chacun a reçu, pour chaque échantillon, une feuille à compléter. Ce questionnaire se composait de deux parties. Dans la première, le jury était invité à cocher son choix concernant le type d'oscillations entendues (vibrato, son blanc, trémolo, chevrottement, ou autre [Figure 30 A]). Dans la deuxième partie, le jury devait se prononcer sur une échelle à cinq points (0-4) pour différents critères [Figure 30 B].

Cochez la case correspondante.

- trémolo	<input type="radio"/>
- vibrato	<input type="radio"/>
- son blanc	<input type="radio"/>
- chevrottement	<input type="radio"/>
- autre	<input type="radio"/> ,

Figure 30 A : Première partie du questionnaire proposée au jury.

Situez sur l'échelle le niveau correspondant (0-4).

Paramètres	0	1	2	3	4	
Non gênant						Gênant
Harmonieux						Dysharmonieux
Agréable						Désagréable
Régulier						Irrégulier
Mature						Embryonnaire
Précis						Imprécis
Stable						Instable
Timbré						Détimbré

Figure 30 B : Deuxième partie du questionnaire proposé au jury.

Nous avons fait l'hypothèse que certains paramètres objectifs du vibrato pouvaient être en corrélation avec l'une ou l'autre caractéristique perceptive. Nous avons choisi en collaboration avec des chanteurs professionnels huit caractéristiques perceptives qui nous semblaient être les plus pertinentes dans la qualification du vibrato. (L'étude 1 précise la procédure de choix des paramètres, p. 105) Nous avons attribué à chacune des ces caractéristiques son opposé, soit " non-gênant/gênant", "harmonieux/dysharmonieux", ... L'échelle comporte cinq points, afin de permettre au jury d'écoute de coter de façon nuancée.

Pour analyser nos échantillons vocaux et les mettre en corrélation avec les paramètres du jugement perceptif, nous avons donc utilisé le Multi-Dimensional Voice Program (M.D.V.P.) modèle 4305 (Kay Elemetrics Corp., Pine Brook, New-Jersey, USA^[159]) détaillé ci-dessus. Ce logiciel permet l'analyse de pathologies de la voix parlée (maladie de Parkinson, dysphonie dysfonctionnelle ou organique, presbyphonie, etc...). Par ailleurs, des modulations périodiques de la fréquence fondamentale (tremor). Il calcule la fréquence et l'étendue des modulations de fréquence et d'amplitude de la fréquence fondamentale (Fo).

Ces mesures nous intéressaient pour l'analyse des trémulations de la voix chantée. Nous n'avons apporté aucun changement à ce logiciel. Nous avons simplement élargi la plage de fréquence de la fenêtre B (oscillogramme de fréquence) dans le cas des chanteuses et des falsettistes afin de capter leur fréquence fondamentale. En effet le programme de base analyse de 30 à 400 Hz mais il est conçu pour permettre une manipulation simple afin de jouer sur cette ordonnée. Nous avons constaté que le M.D.V.P n'analyse pas les fréquences supérieures à ré4 soit 587 Hz. Ceci explique l'absence de données pour un sujet féminin dans la tâche du [a] vibré. Nous avons utilisé un D.A.T. 20 (Digital Audio Tape) (Teac Corp. ; Nakasho, Musashino-Shi, Tokyo, Japon), et un micro A.K.G. C 1000 S (Vienne, Autriche) pour capter nos échantillons vocaux.

Les enregistrements ont été réalisés, pour la plupart, au domicile du chanteur. Nous avons été attentifs au choix d'un endroit calme. Le micro sur pied, détaillé ci-dessus, était situé à trente centimètres du chanteur et ceci pendant les quatre tâches.

L'ensemble du système informatique pouvait capter les données digitales du D.A.T. et analyser les échantillons recueillis.

Nous avons choisi six paramètres chiffrés du M.D.V.P. En référence aux paramètres utilisés dans la littérature, il nous a paru essentiel de tenir compte de l'étendue et de la cadence des variations de fréquence et d'amplitude du vibrato (ou autres trémulations), ainsi que du jitter qui correspond à l'irrégularité de la variation de fréquence et du shimmer qui exprime l'irrégularité de la variation d'amplitude. Ces deux mesures reflètent les variations de modèles vibratoires des cordes vocales et leur instabilité.

Les 6 paramètres sélectionnées sont :

- **Fftr** (Fundamental-Tremor Frequency) qui exprime en Hertz **la cadence de variation de fréquence** ;
- **Fatri** (Amplitude Tremor Frequency) qui exprime en Hertz **la cadence de variation d'amplitude** ;
- **Ftri** (Fundamental Tremor Intensity Index) qui exprime en pourcentage **l'étendue de la variation de fréquence**, donnée représentée en ordonnée dans la fenêtre D [annexe 9].
- **Atri** (Amplitude Tremor Intensity Index) qui exprime en pourcentage **l'étendue de la variation d'amplitude**, donnée représentée en ordonnée dans la fenêtre E [annexe 9].
- **Jitt** (Jitter Percent) qui représente **l'évaluation en pourcentage de la variabilité**, période par période (à court terme), **de la hauteur tonale** de la vocalisation ;
- **Shim** (Shimmer Percent) qui représente **l'évaluation en pourcentage de la variabilité**, période à période (à court terme), **de l'amplitude** dans l'échantillon vocal.

Nous avons retenu ces six paramètres pour chaque [a] vibré, [a] non vibré et pour une voyelle du morceau imposé : [a], il s'agit du [an] de temps dans la quatrième mesure. Nous avons prélevé la dernière seconde et demi de cette mesure. Le chanteur y produisait généralement un vibrato, juste après un trille.

5.3.3. Résultats.

- Fidélité intra-juges.

Pour la **concordance de l'observateur par rapport à lui-même**, nous avons voulu savoir, dans un premier temps, si le jugement de chaque membre était constant (Fidélité test/re-test). Nous avons calculé le pourcentage de réponses concordantes après les deux écoutes (Percent Agreement). Seul le juge 1 est concordant et ce pour les quatre tâches soit le [a] vibré 72% de concordance, le [a] non vibré 87% de concordance, le morceau choisi (M.C.) 73% de concordance, et le morceau imposé (M.I.) 67% de concordance [Tableau 55].

Tableau 55: Concordance du juge par rapport à lui-même par tâche concernant le type d'oscillation.

TACHES	résultats concordants	
	n	%
M.C. (n=30)	22	73
	21	70
	9	30
M.I. (n=30)	20	67
	14	47
	9	30
a vibré (n=29)	21	72
	13	45
	8	28
a non- vibré (n=30)	26	87
	22	73
	17	57

Légende: M.C.= morceau choisi ; M.I.= morceau imposé ; N : nombre de chanteurs ayant effectué la tâche.

Comme l'indique le tableau 56. nous constatons que le juge 1 est significativement différent du juge 2 pour le [a] vibré ($p=0.033$). Il l'est aussi significativement du juge 3 pour toutes les tâches, [a] vibré ($p<10^{-3}$), [a] non-vibré ($p=0.01$), M.C. ($p<10^{-3}$), M.I. ($p=0.004$).

Le juge 2 est significativement différent du juge 3 pour le M.C. ($p=0.002$).

Tableau 56 : **Comparaison des pourcentages d'observations concordantes par tâche concernant le type d'oscillation.**

Tâches	Juge1 vs 2	Juge 1 vs 3	Juge 2 vs 3
M.C.	NS	$p < 10^{-3}$	$p = 0.002$
M.I.	NS	$p = 0.004$	NS
a vibré	$p = 0.033$	$p < 10^{-3}$	NS
a non-vibré	NS	$p = 0.01$	NS

Légende: M.C.= morceau choisi ; M.I.= morceau imposé

Dans un deuxième temps, nous avons voulu savoir qu'elle était la tâche qui obtenait la meilleure concordance, tous les juges confondus. La concordance semble meilleure pour le [a] non vibré, mais avec un index Kappa²⁷ (indice de concordance entre deux juges) peu élevé ($p = 0.017$, $K = 0.307$) [Tableau 57].

Tableau 57: **Concordances du juge par rapport à lui-même par tâche concernant le type d'oscillation.**

Tâches	Résultats concordants				
	N	%	Kappa	SE (K)	P
M.C.	52	58	0.421	0.071	$p < 10^{-5}$
M.I.	43	48	0.312	0.069	$p < 10^{-5}$
a vibré (n= 87)	42	48	0.324	0.07	$p < 10^{-5}$
a non-vibré	65	72	0.307	0.118	$p = 0.017$

Légende: M.C.= morceau choisi ; M.I.= morceau imposé.

Nous avons également cherché quel était le juge le plus concordant, toutes tâches confondues. Le juge 1 est significativement plus concordant que le juge 2 ($p = 0.009$), lequel est plus concordant que le juge 3 ($p < 10^{-3}$). [Tableau 58 et 59].

²⁷ **L'index Kappa** : Il est égal au nombre de concordances observées moins le nombre de concordances dues au hasard, rapporté à l'ensemble des concordances possibles en-dehors de la chance. Il varie de 0 (concordance par chance) à 1 (concordance parfaite). D'après Landis & Koch, un index de 0.8 équivaut à une bonne concordance.

Tableau 58: **Concordances des juges par rapport à eux-mêmes par juge concernant le type d'oscillation.**

Jury	N	%	Kappa	SE (K)	P
1	89	75	0.635	0.058	$p < 10^{-5}$
2	70	59	0.419	0.064	$p < 10^{-5}$
3	43	36	0.18	0.057	$p < 10^{-3}$

Tableau 59: **Comparaison des pourcentages d'observations concordantes par juge concernant le type d'oscillation.**

	Compar.% obs. conc.(X ²).	Compar. Kappa (Z)
1 vs 2	p= 0.009	p= 0.012
1 vs 3	$p < 10^{-5}$	$p < 10^{-5}$
2 vs 3	$p < 10^{-3}$	p= 0.005

Nous avons voulu déterminer quelle était la corrélation entre les deux lectures, des huit paramètres de l'échelle pour chaque membre du jury dans chacune des quatre tâches [Tableau 60]. L'ensemble des coefficients de Spearman²⁸ a été comparé entre les membres du jury par un test de Wilcoxon²⁹. Les juges 2 et 3 ne diffèrent pas. Le juge 1 n'a pas été comparé vu le nombre d'abstentions. En effet, le juge 1 avait choisi de ne pas répondre aux paramètres de régularité, de maturité, de précision et de stabilité lorsqu'il avait à porter un jugement sur un son blanc.

Nous avons donc recalculé un coefficient de Wilcoxon en considérant les quatre variables pour lesquelles le juge 1 avait participé: "non-gênant/gênant", "harmonieux/dysharmonieux", "agréable/désagréable", "timbré/détimbré". Le juge 1 est significativement meilleur (0.71 vs 0.388 et 0.278). [Tableau 60].

²⁸ **Le coefficient de Spearman** : coefficient de corrélation lorsque les deux ensembles de valeurs sont remplacés par leurs rangs (de 1 à n).

²⁹ **Le test de Wilcoxon** utilise les rangs des valeurs observées, il est indépendant de leur distribution.

Tableau 60 : Corrélations des juges par rapport à eux-mêmes par juge pour 4 variables et 4 tâches pour la question 2 concernant les 8 paramètres.

JURY	Quartile 1	Médiane	Quartile 3
1	0.614	0.71	0.832
2	0.228	0.388	0.502
3	0.156	0.278	0.526

En considérant les quatre variables précédentes, nous avons déterminé le degré de concordance par tâche, les trois juges confondus. La concordance est meilleure pour le morceau choisi (M.C.) par rapport aux trois autres tâches (M.C. vs M.I.: $p=0.015$; M.C. vs [a] vibré $p=0.028$; M.C. vs [a] non vibré: $p=0.002$). Elle est intermédiaire pour le morceau imposé (M.I.) et le [a] vibré. Elle est mauvaise pour le [a] non- vibré (MI vs [a].vibré: NS ; MI vs [a] non vibré: $p=0.004$; [a] vibré vs [a] non vibré: $p=0.008$) [Tableau 61].

Tableau 61: Comparaison intra-juges 2 à 2 par tâche, en considérant 4 variables pour la question 2 concernant les 8 paramètres.

Tâches	P
M.C. vs M.I.	$p= 0.015$
M.C. vs [a] vibré	$p= 0.028$
M.C. vs [a] non-vibré	$p= 0.002$
M.I. vs [a] vibré	N.S.
M.I. vs [a] non-vibré	$p= 0.004$
A vibré vs [a] non-vibré	$p= 0.008$

Légende: M.C.= morceau choisi ; M.I.= morceau imposé

- Fidélité inter-juges.

Pour la **concordance des observateurs entre eux**, les lectures des juges 2 et 3 étant peu reproductibles, il semble sans intérêt de les comparer à celles du juge 1.

- Corrélation paramètres objectifs/paramètres subjectifs.

En ce qui concerne **les corrélations entre paramètres objectifs et paramètres subjectifs**, nous avons d'abord cherché les relations entre, d'une part les réponses du jury 1. et d'autre part les 6 paramètres objectifs du M.D.V.P., **Fftr** (Fo Tremor Frequency), **Fatr** (Amplitude Tremor Frequency), **Ftri** (Fo Tremor Intensity Index), **Atri** (Amplitude Tremor Intensity Index), **Jitt** (Jitter percent), **Shim** (Shimmer percent). Nous avons considéré les seules réponses où le juge 1 était concordant, c'est-à-dire les deux catégories vibrato et son blanc, pour les tâches [a] vibré et [a] (morceau imposé). Le test de Wilcoxon nous donne les résultats suivants, dans le cas du M.I., le **Ftri** et le **Jitt** sont significativement différents ($p=0.003$ vs $p=0.017$) entre vibrato et son blanc. **Ftri** et **Jitt** sont plus élevés dans le cas du vibrato et plus bas dans le cas du son blanc. Dans le cas du [a] vibré, le **Ftri** et le **Shim** sont significativement différents ($p<10^{-3}$ et $p=0.033$) entre vibrato et son blanc. **Ftri** et **Shim** sont élevés dans le cas du vibrato et plus bas pour le son blanc [Tableau 62].

Tableau 62: **Corrélations entre paramètres subjectifs et objectifs concernant le type d'oscillation et le juge 1.**

Tâches	Param.	Vibrato			Son blanc			Wilcoxon
		N=9			N=7			Echantillon
		Quartile 1	Médiane	Quartile 3	Quartile 1	Médiane	Quartile 3	Indép. Comparaison
M.I.	Fftr	2.08	5.97	6.35	1.72	5.48	6.56	NS
M.I.	Fatr	4.71	5.48	6.35	1.62	5.8	6.45	NS
M.I.	Ftri	2.63	3.93	4.67	0.47	0.82	1.05	P= 0.003
M.I.	Atri	9.12	12.87	16.42	5.66	7.62	13.99	NS
M.I.	Jitt	1.35	1.53	2.03	0.34	0.77	1.38	P= 0.017
M.I.	Shim	5	6.41	7.49	1.45	2.42	6.77	P= 0.081
			N=8			N=10		
[a] V	Fftr	5.2	5.41	5.76	4.76	5.17	6.25	NS
[a] V	Fatr	5.41	5.9	7.05	3.31	5.1	6.45	NS
[a] V	Ftri	1.33	1.88	2.33	0.31	0.4	0.51	$P< 10^{-3}$
[a] V	Atri	9.86	12.82	15.41	5.84	10.6	13.41	NS
[a] V	Jitt	0.36	0.4	0.7	0.23	0.39	0.51	NS
[a] V	Shim	2.88	3.18	3.91	1.7	2.07	2.25	P= 0.033
[a] N-V						n=24		
					Pas de comparaison possible.			

Légende: [a] V= [a] vibré ; [a] NV: [a] non vibré.

Ensuite, nous avons cherché les corrélations entre, d'une part les réponses du juge 1 aux quatre paramètres "non-gênant/gênant", "harmonieux/dysharmonieux", "agréable/désagréable", "timbré/détimbré", pour les trois tâches [a] vibré, [a] non vibré et M.I. (morceau imposé), et d'autre part, les six paramètres objectifs fournis par le M.D.V.P.. Nous avons trouvé les corrélations suivantes, **Ftri** est significativement corrélé pour M.I. et [a] vibré, dans les quatre paramètres. **Jitt** est significativement corrélé pour le M.I. uniquement, dans les quatre paramètres. **Shim** est significativement corrélé pour le M.I. et le [a] vibré dans le paramètre timbré, et dans le [a] non vibré dans le paramètre harmonieux. Chacun de ces paramètres (**Ftri**, **Jitt**, et **Shim**) est négativement corrélé aux tâches précitées, c'est-à-dire que le juge 1 a coté d'autant plus haut sur l'échelle de niveau (0-4) que les paramètres objectifs sont bas [Tableau 63].

Tableau 63: Corrélations entre paramètres subjectifs et objectifs concernant les paramètres et le juge 1.

Paramètres objectifs	Paramètres subjectifs	Morceau Imposé.		A vibré		a non-vibré	
		Rs	Signif	Rs	signif	Rs	Signif
Fftr	Non-gênant	0.269	NS	0.224	NS	0.175	NS
	Harmonieux	0.177	NS	0.163	NS	0.22	NS
	Agréable	0.195	NS	0.171	NS	0.215	NS
	Timbré	0.34	NS	0.232	NS	0.155	NS
Fatr	Non-gênant	0.046	NS	0.258	NS	0.029	NS
	Harmonieux	0.082	NS	0.203	NS	0.012	NS
	Agréable	0.111	NS	0.205	NS	0.026	NS
	Timbré	0.143	NS	0.428	0.023	0.132	NS
Ftri	Non-gênant	0.519	0.004	0.449	0.015	0.198	NS
	Harmonieux	0.526	0.003	0.54	0.002	0.21	NS
	Agréable	0.547	0.002	0.543	0.002	0.197	NS
	Timbré	0.606	<10⁻³	0.548	0.003	0.309	NS
Atri	Non-gênant	0.067	NS	0.123	NS	0.013	NS
	Harmonieux	0.099	NS	0.144	NS	0.032	NS
	Agréable	0.099	NS	0.145	NS	0.045	NS
	Timbré	0.27	NS	0.302	NS	0.273	NS
Jitt	Non-gênant	0.397	0.033	0.049	NS	0.202	NS
	Harmonieux	0.443	0.016	0.006	NS	0.222	NS
	Agréable	0.399	0.032	0	NS	0.136	NS
	Timbré	0.375	0.045	0.085	NS	0.212	NS
Shim	Non-gênant	0.22	NS	0.281	NS	0.355	0.059
	Harmonieux	0.276	NS	0.357	0.057	0.394	0.034
	Agréable	0.25	NS	0.355	0.059	0.297	NS
	Timbré	0.407	0.028	0.43	0.022	0.04	NS

5.3.4. Discussion et conclusion.

Par rapport à la première partie de la grille de perception [Figure 30 A] les résultats de la **concordance intra-observateurs** mettent en évidence qu'un seul membre du jury est fidèle au test/re-test. Nous expliquons ce résultat par le fait que de nombreux facteurs ont pu intervenir : des facteurs propres au juge, tels que l'expérience professionnelle, les connaissances scientifiques, le recul par rapport à la profession ; des facteurs extérieurs au sujet tels que le manque de temps, la disponibilité, le stress de la contrainte professionnelle, l'impossibilité de se focaliser sur la trémulation sans tenir compte du timbre de la voix, de l'interprétation, de la tessiture, du style de musique, des connaissances en psycho-acoustique, etc.

Les résultats nous suggèrent également que le type de tâche n'aurait pas d'influence sur la reproductibilité du jugement, autrement dit sur un [a] tenu, le juge n'aurait pas plus de facilité à émettre un avis constant, bien qu'il s'agissait d'un échantillon court et que le juge aurait pu être moins influencé par le type de musique, l'interprétation ainsi que la grande variabilité de fréquence et d'intensité.

Par rapport à la deuxième partie de la grille de perception [tableau 30 B] les résultats montrent ensuite que le juge 1 reste significativement concordant pour 4 paramètres "non-gênant/gênant", "harmonieux/dysharmonieux", "agréable/désagréable", "timbré/détimbré". Le fait qu'un seul membre du jury d'écoute ait montré une constance de jugement, met en question la fiabilité de l'oreille. Les résultats mettent en évidence également un effet du type de tâche. Dans l'ensemble, le jury est plus constant dans ses cotations pour les morceaux choisis. Nous supposons l'hypothèse que la longueur de l'échantillon aurait pu jouer sur la constance de la cotation. En une vingtaine de mesures, le jury avait le temps de se fixer des repères auditifs. Il avait aussi plus de temps pour se prononcer sur les paramètres subjectifs ce qui pourrait jouer sur la finesse de son jugement.

La question de la **concordance entre les juges** ne se pose pas, dans la mesure où il faudrait au minimum deux sujets à comparer.

Pour la **corrélation entre paramètres subjectifs et objectifs**, il est intéressant de noter que les réponses du juge 1 sont significativement corrélées avec les

paramètres objectifs **Ftri**, **Jitt**, et **Shim** du M.D.V.P.. L'oreille humaine, avec le talent, l'entraînement, l'expérience, et probablement d'autres facteurs, peut donc être sensible à certains paramètres vocaux. Cette extraordinaire capacité de l'oreille semble être une exception.

Nous attirons l'attention sur le fait que le M.D.V.P. est conçu pour l'analyse de la voix parlée. Nous supposons que les mesures de tremor de la voix chantée sont insuffisantes de par l'algorithme de calcul du programme. Par exemple le **Fftr** qui mesure la cadence de tremor, est sensible au paramètre d'irrégularité. Nous avons constaté avec l'aide de physiciens acousticiens, que les données chiffrées du **Fftr** n'étaient fiables que dans le cas où l'oscillation était régulière pendant toute la durée du signal. Il serait intéressant d'étudier la corrélation entre le **Fftr** dans le cas d'oscillations régulières, et les paramètres subjectifs du juge 1.

Nous pensons que la cadence d'oscillation de fréquence (**Fftr**) et d'amplitude (**Fatr**) pourraient constituer des paramètres déterminants. La cadence d'oscillation de fréquence a été plusieurs fois citée dans la littérature (Miller (1986)^[200] ; Reid (1972)^[246] ; Hakes et al (1988)^[107] ; Sapir et al (1993)^[266] ; Cleveland (1994)^[38]). Les cadences avancées par ces auteurs se groupaient entre 5 et 7 pulsations/sec. Rothman et al (1990)^[261] avaient constaté que ce paramètre recevait une forte proportion de jugements "bon vibrato", dans une fourchette équivalente: 5.1 à 6.9 pulsations/sec.

Quant à la cadence d'oscillation d'amplitude, peu citée dans la littérature, Rothman et al (1990)^[261] avaient mis en évidence qu'elle était corrélée avec les jugements "bon vibrato" dans une plage de fréquence de 5 à 6 pulsations/sec. Nous comptons sur ces deux paramètres (**Fftr** et **Fatr**) mais les limites de leur algorithme de calcul au M.D.V.P. ne nous pas permis d'en tenir compte.

Nous pouvons conclure que l'écoute de l'oreille humaine est supérieure aux possibilités des logiciels d'analyse de la voix pour le juge 1. qui peut être considéré comme une oreille exceptionnelle. Ceci plaide donc en faveur de l'utilisation de logiciels apportant des données fiables.

Cette étude a montré que le paramètre d'étendue de l'oscillation de fréquence (**Ftri**) est pertinent dans l'étude du vibrato, ainsi que le Jitter (**Jitt**) et le Shimmer (**Shim**). Les

paramètres de cadence d'oscillation de fréquence (**Fftr**) et d'amplitude (**Fatr**) seraient probablement déterminants s'ils étaient adaptés par des spécialistes à l'étude de la voix chantée.

5.4. ÉTUDE 5 : APPLICATION DE LA STATION E.V.A. 1 À UNE POPULATION D'ENFANTS ÂGÉS DE 4 À 9 ANS CHANTEURS ET NON CHANTEURS.

5.4.1. Introduction.

L'évaluation objective de la qualité vocale a fait l'objet de nombreuses études. Celles-ci visent à confirmer son efficacité, à en améliorer la pertinence, la précision et la fiabilité. Les méthodes d'investigations à partir des logiciels sont en plein essor, elles affinent l'analyse de la voix pour les besoins de la recherche et de la pratique. Néanmoins les valeurs de référence font défauts. Nous éprouvons donc des difficultés à comparer les résultats obtenus de nos sujets à un "standard". Cette étude a pour but de contribuer à l'établissement de valeurs de référence pour une population d'enfants âgés de 4 à 9 ans. Afin de les éprouver, nous avons voulu mesurer quel serait l'effet d'un entraînement vocal sur les mesures acoustiques et aérodynamiques fournies par la station E.V.A.1.

5.4.2. Matériel et méthode.

Notre échantillon est constitué de 103 enfants âgés de 4 à 9 ans, répartis en deux groupes de 51 filles et 52 garçons, issus des écoles communales de Visé et d'Yvoir. Les enfants pour lesquels une pathologie vocale objectivée nous a été communiquée, par l'intermédiaire des parents ou de l'enseignant, se sont vus écartés. Nous avons l'accord de l'enfant et de ses parents pour une contribution anonyme au corpus. Nous leur avons demandé d'émettre un [a] le plus longtemps possible une première fois en voix parlée et une deuxième fois en voix chantée. La hauteur tonale et l'intensité de l'émission sonore sont laissées à l'initiative de l'enfant. Malgré le fait que l'appareil ne nécessite pas de cabine anéchoïde, nous avons demandé à réaliser les examens dans un endroit calme et isolé. Aucune consigne précise n'a été

préalablement donnée aux enfants. Les enfants sont entrés dans le local 3 par 3 ce qui a permis d'atténuer la timidité de certains. Notre souci est d'éviter tout stress. Les consignes, explications et familiarisation sont abordées à cet instant. Nous évitons de faire entendre un exemple afin de laisser à l'enfant le choix de sa tonalité. L'enfant est debout face à l'ordinateur. Il est invité à placer hermétiquement sa bouche dans la pièce à main en silicone souple [cfr figure 24 B]. Nous veillons à régler la hauteur du pied parfaitement à sa hauteur et nous assurons que l'embouchure s'accolle au visage de l'enfant. L'ensemble de la prise d'informations ne prend pas plus de 10 minutes par enfant. Les essais se limitent à trois émissions. L'examineur retient l'émission qui lui semble correspondre à la voix spontanée du sujet. Nous n'avons noté aucun refus de participation de la part des enfants. Ils sont motivés, attentifs et respectent le silence imposé durant la tâche. La tâche se veut simple, courte et rapide.

Notre analyse s'est portée d'une part sur la partie jugée la plus stable de l'émission sonore et d'autre part, sur le début de l'émission sonore. Les paramètres acoustiques et aérodynamiques sont calculés simultanément.

Le même design a été appliqué à une population de 86 enfants chanteurs issus du groupe des Cantourelles³⁰ du mouvement international "A cœur joie" dans le but de les comparer à nos valeurs de référence établies sur des sujets sans entraînement vocal. Les choristes répètent à raison de trois quart d'heure par semaine en moyenne et mettent en place un répertoire populaire qu'ils présentent à l'occasion de concert. Le chef de chœur propose un travail de sensibilisation au niveau :

- de la posture pour la stabilité en position debout,
- du rythme,
- de l'articulation en mettant l'accent sur l'aperture buccale et la diction,
- de la projection du son pour la maîtrise de l'intensité,
- de l'écoute mutuelle pour harmoniser les timbres.

³⁰ Les **cantourelles** : le groupement chorale "A Cœur joie" répartit les enfants en deux niveaux : les cantourelle pour les 4 à 9 ans et la chanterie pour les 9-12 ans. Les Cantilènes sont réservées aux adolescents.

5.4.3. Résultats.

a. Statistiques descriptives pour les enfants sains.

- Pour la tâche parlée

Tableau 64 A : Valeurs obtenues pour le début et la partie stable de l'émission.

[a] parlé	Début de l'émission (attaque incluse)							
Mesures	Moyenne		DS		Minimum		Maximum	
sexes	F	G	F	G	F	G	F	G
Fo moyen	271	288	42.82	42.18	199	209	412	466
Fo écart	18.55	19.58	11.35	11.16	3.00	5.00	59.00	46.00
Fo %	7.02	6.82	4.76	3.92	1.40	1.90	23.50	15.90
Int moyen	79.20	75.93	5.70	6.77	62.32	56.78	90.73	85.79
Int écart	4.25	4.02	2.12	1.73	0.88	0.76	8.51	7.68
Int %	5.39	5.27	2.41	2.16	1.10	1.10	9.60	9.60
DAB moy	2.55*	-1.07*	7.90*	8.55*	-0.18	-0.17	0.15	0.14
DAB écart	1.54*	1.27*	8.13**	8.15**	0.00	0.00	0.04	0.03
DAB %	13.26	8.33	22.04	45.05	-49.30	-242.20	49.00	107.20
Jitter sdt	5.15	5.04	5.05	4.10	1.70	1.90	29.90	20.00
Jitter %	1.95	1.79	2.05	1.45	0.60	0.40	11.80	7.10
Shim sdt	0.18	0.16	5.77*	5.25*	0.10	0.10	0.40	0.30
Shim %	0.22	0.21	7.37*	6.15*	0.10	0.10	0.50	0.30
ShimD sdt	0.84	0.70	0.28	0.26	0.43	0.09	1.77	1.34
ShimD %	0.85	1.33	1.12	5.89	-1.70	-9.30	2.90	38.40
Fuite Glo	4.81	4.52	1.83	1.83	1.94	2.07	9.12	11.99
[a] parlé	Partie stable de l'émission							
Mesures	Moyenne		DS		Minimum		Maximum	
Sexes	F	G	F	G	F	G	F	G
Fo moyen	271	286	44.53	42.93	203	208	422	465
Fo écart	4.12	6.02	3.37	14.80	1.00	1.00	19.00	106.00
Fo %	1.52	2.26	1.32	6.18	0.50	0.60	7.90	44.50
Int moyen	78.96	75.31	6.37	6.51	62.26	58.54	90.94	85.56
Int écart	0.75	0.79	0.46	0.43	0.25	0.28	3.04	2.33
Int %	0.97	1.07	0.66	0.62	0.30	0.40	4.30	2.90
DAB moy	2.14*	-1.60*	8.03*	8.55*	-0.18	-0.17	0.15	0.14
DAB écart	6.35**	4.42**	3.82**	2.93**	0.00	0.00	0.02	0.01
DAB %	6.69	9.38	9.29	24.54	-16.30	-13.50	23.50	164.80
Jitter sdt	1.56	2.13	1.36	2.74	0.50	0.70	7.80	19.90
Jitter %	0.57	0.87	0.53	1.31	0.20	0.30	3.20	7.20
Shim sdt	9.22*	7.12*	5.23*	4.99*	0.00	0.00	0.20	0.20
Shim %	0.12	9.81*	5.79*	5.05*	0.00	0.00	0.30	0.20
ShimD sdt	0.55	0.50	0.13	0.13	0.34	0.28	0.89	0.92
ShimD %	0.67	2.10	0.99	9.58	-1.10	-4.60	4.60	68.40
Fuite Glo	7.38	7.40	3.36	3.66	2.45	0.97	15.64	21.58

Légende : *: valeurs x 100 ; ** valeurs x 1000.

- Pour la tâche chantée

Tableau 64 B : Valeurs obtenues pour le début et la partie stable de l'émission.

[a] chanté	Début de l'émission (attaque incluse)							
Mesures	Moyenne		DS		Minimum		Maximum	
Sexes	F	G	F	G	F	G	F	G
Fo moyen	320	335	91.54	100.68	184	221	660	690
Fo écart	21.31	17.38	14.99	11.91	5.57	3.00	83.00	55.00
Fo %	6.95	4.94	4.62	3.23	1.90	-3.80	31.30	18.10
Int moyen	81.58	75.63	6.21	8.59	69.54	50.83	94.00	90.94
Int écart	4.60	4.55	2.19	2.92	0.76	1.02	8.43	15.00
Int %	5.61	5.72	2.56	3.39	1.00	1.20	11.00	21.10
DAB moy	3.28*	-5.00*	8.32*	9.21*	-0.17	-0.16	0.19	0.30
DAB écart	1.74*	1.80*	1.08*	2.86*	0.00	0.00	0.06	0.20
DAB %	21.11	3.14	50.40	71.75	-74.60	-456.40	306.00	118.50
Jitter sdt	5.67	4.52	7.42	5.63	1.30	1.00	43.40	36.70
Jitter %	1.91	1.39	2.57	1.74	0.20	0.40	16.40	12.00
Shim sdt	0.15	0.14	5.05*	5.02*	0.10	0.10	0.20	0.20
Shim %	0.18	0.19	5.55	8.02	0.10	0.10	0.30	0.50
ShimD sdt	0.87	0.72	0.41	0.45	0.26	0.30	1.91	3.00
ShimD %	1.55	0.96	3.20	2.49	-2.00	-4.80	19.40	13.50
Fuite Glo	5.95	5.13	2.91	3.57	2.21	1.85	14.56	24.66
[a] chanté	Partie stable de l'émission							
Mesures	Moyenne		DS		Minimum		Maximum	
Sexes	F	G	F	G	F	G	F	G
Fo moyen	317	337	89.40	98.92	185	213	615	677
Fo écart	6.29	4.69	6.56	4.02	0.00	1.00	35.00	21.00
Fo %	2.12	1.46	2.65	1.53	0.00	0.50	15.20	9.80
Int moyen	81.29	75.12	7.04	8.30	68.76	55.53	95.50	91.16
Int écart	1.08	1.35	1.19	2.40	0.29	0.25	7.63	16.27
Int %	1.36	1.89	1.64	3.58	0.40	0.40	10.80	24.10
DAB moy	2.86*	-8.50*	8.63*	8.96*	-0.17	-0.17	0.20	0.21
DAB écart	7.26*	6.39*	5.22*	6.21*	0.00	0.00	0.02	0.04
DAB %	10.32	5.91	16.39	15.61	-21.10	-53.70	72.70	50.20
Jitter sdt	2.47	1.65	3.38	0.86	0.00	0.70	19.80	5.40
Jitter %	0.86	0.52	1.40	0.32	0.00	0.20	8.60	2.00
Shim sdt	6.86*	7.12*	5.47*	6.05*	0.00	0.00	0.20	0.20
Shim %	8.24*	9.81*	5.18*	8.04*	0.00	0.00	0.20	0.30
ShimD sdt	0.55	0.48	0.22	0.18	0.00	0.26	1.09	1.17
ShimD %	0.90	0.89	1.76	2.58	-4.00	-2.40	8.20	17.40
Fuite Glo	9.12	8.59	5.09	5.30	0.00	1.79	27.90	33.87

Légende : *: valeurs x 100 ; ** valeurs x 1000.

b. Statistiques descriptives des enfants chanteurs.

- Pour la tâche parlée.

Tableau 64 C : Valeurs obtenues pour le début et la partie stable de l'émission.

[a] parlé	Début de l'émission (attaque incluse)							
Mesures	Moyenne		DS		Minimum		Maximum	
Sexes	F	G	F	G	F	G	F	G
Fo moyen	268	272	26.92	28.65	213	214	320	379
Fo écart	15.08	15.21	4.62	7.66	4.00	3.00	22.00	35.00
Fo %	5.70	5.66	1.71	2.98	1.70	1.00	8.60	14.10
Int moyen	73.96	71.53	5.25	4.99	64.42	54.42	87.52	83.21
Int écart	5.21	4.94	1.38	1.47	2.15	1.88	8.07	11.76
Int %	7.02	6.93	1.67	2.37	3.00	2.80	11.20	21.60
DAB moy	7.79*	5.97*	4.73*	3.79*	0.00	0.00	0.17	0.23
DAB écart	1.90*	1.47*	1.14*	1.13*	0.00	0.00	0.04	0.07
DAB %	29.59	25.30	25.53	21.05	10.40	-89.10	148.90	76.60
Jitter sdt	3.24	3.38	1.56	2.24	1.30	1.00	7.50	16.70
Jitter %	1.19	1.26	0.48	0.85	0.50	0.40	2.40	6.10
Shim sdt	0.12	0.14	4.36*	5.59*	0.10	0.10	0.20	0.30
Shim %	0.21	0.21	4.40*	7.85*	0.10	0.10	0.30	0.40
ShimD sdt	0.76	0.68	0.27	0.20	0.39	0.34	1.58	1.47
ShimD %	1.57	1.64	1.71	3.73	0.60	-16.60	8.90	21.40
Fuite Glo	5.33	4.91	1.77	1.67	2.35	1.53	9.01	10.04
[a] parlé	Partie stable de l'émission							
Mesures	Moyenne		DS		Minimum		Maximum	
Sexes	F	G	F	G	F	G	F	G
Fo moyen	266.48	271.56	26.70	31.54	215	209	326	398
Fo écart	2.68	2.82	1.14	1.54	1.00	1.00	6.00	11.00
Fo %	1.04	1.05	0.40	0.49	0.50	0.50	2.10	3.40
Int moyen	72.63	71.50	5.39	4.82	63.40	59.86	87.67	83.53
Int écart	0.63	0.82	0.34	0.68	0.12	0.19	1.39	3.77
Int %	0.87	1.15	0.48	0.98	0.20	0.20	2.00	5.80
DAB moy	6.20*	5.23*	4.22*	3.60*	-0.04	-0.01	0.14	0.20
DAB écart	5.36**	5.08**	4.18**	6.48**	0.00	0.00	0.02	0.05
DAB %	6.32	11.82	17.02	12.35	-71.40	-31.30	21.20	63.20
Jitter sdt	1.16	1.20	0.35	0.40	0.70	0.70	2.00	2.40
Jitter %	0.43	0.44	0.13	0.14	0.20	0.30	0.80	1.00
Shim sdt	2.00*	3.44*	5.00*	4.79*	0.00	0.00	0.20	0.10
Shim %	5.20*	6.23*	7.14*	6.37*	0.00	0.00	0.20	0.20
ShimD sdt	0.47	0.44	0.14	0.12	0.28	0.25	0.79	0.93
ShimD %	0.60	1.62	2.16	3.09	-9.20	-6.20	3.00	20.30
Fuite Glo	13.91	11.16	7.09	6.29	2.18	2.96	30.28	26.92

Légende : * : valeurs x 100 ; ** valeurs x 1000.

- Pour la tâche chantée

Tableau 64 D : Valeurs obtenues pour le début et la partie stable de l'émission.

[a] chanté	Début de l'émission (attaque incluse)							
Mesures	Moyenne		DS		Minimum		Maximum	
Sexes	F	G	F	G	F	G	F	G
Fo moyen	318	318	52.84	90.33	260	226	504	768
Fo écart	16.80	16.57	7.62	8.41	5.00	3.00	34.00	42.00
Fo %	5.24	5.22	2.22	2.33	1.50	1.00	9.40	11.50
Int moyen	73.94	71.67	5.39	5.52	61.93	59.43	84.02	81.30
Int écart	5.74	4.98	1.32	1.60	3.15	2.08	8.33	8.52
Int %	7.73	6.90	1.55	1.99	4.30	3.30	10.30	11.90
DAB moy	6.57*	6.00*	4.89*	3.86*	0.00	0.00	0.18	0.17
DAB écart	1.73*	1.31*	1.41*	7.58**	0.00	0.00	0.05	0.04
DAB %	11.71	30.09	71.49	32.83	-328.00	8.60	54.70	256.60
Jitter sdt	3.64	3.29	1.67	1.61	1.10	1.10	9.60	9.20
Jitter %	1.16	1.08	0.58	0.54	0.40	0.30	3.10	3.10
Shim sdt	0.14	0.14	4.90*	5.82*	0.10	0.10	0.20	0.30
Shim %	0.18	0.18	5.77*	8.00*	0.10	0.10	0.30	0.40
ShimD sdt	0.68	0.60	0.38	0.25	0.17	0.20	1.61	1.30
ShimD %	0.45	2.01	4.38	2.78	-20.30	0.20	3.60	17.80
Fuite Glo	5.06	4.63	2.61	1.86	2.01	1.83	11.54	8.79
[a] chanté	Partie stable de l'émission							
Mesures	Moyenne		DS		Minimum		Maximum	
Sexes	F	G	F	G	F	G	F	G
Fo moyen	311	320	56.93	89.16	249	224	512	762
Fo écart	3.20	3.58	1.89	2.62	1.00	1.00	11.00	16.00
Fo %	1.00	1.10	0.38	0.74	0.40	0.50	2.10	5.40
Int moyen	73.38	71.78	6.16	6.26	61.59	59.95	85.58	84.02
Int écart	0.75	0.92	0.45	0.63	0.21	0.20	2.04	0.50
Int %	1.06	1.30	0.69	0.92	0.30	0.30	3.30	5.90
DAB moy	5.65*	5.30*	4.77*	3.83*	-0.01	0.00	0.18	0.16
DAB écart	6.00**	4.92**	5.67**	3.10*	0.00	0.00	0.02	0.01
DAB %	11.58	14.35	9.10	16.68	-14.30	0.12	31.80	117.10
Jitter sdt	1.34	1.67	0.41	1.71	0.80	0.70	2.80	14.30
Jitter %	0.43	0.50	0.12	0.28	0.20	0.30	0.80	2.40
Shim sdt	2.00*	3.44*	4.08*	4.79*	0.00	0.00	0.10	0.10
Shim %	4.00*	6.89*	6.46*	6.47*	0.00	0.00	0.20	0.20
ShimD sdt	0.43	0.41	0.16	0.14	0.15	0.16	0.97	0.88
ShimD %	1.11	1.72	1.82	2.31	-4.80	0.10	5.60	12.20
Fuite Glo	13.19	10.43	7.32	5.69	3.07	2.87	32.57	27.85

Légende : * : valeurs x 100 ; ** valeurs x 1000.

c. Comparaison chanteurs versus non-chanteurs.

Les résultats analysent les différences vocales entre les filles et les garçons selon les mesures acoustiques et aérodynamiques récoltées par le logiciel E.V.A. station 1 .

Nous avons effectué un test d'Aspin-Welch (au test d'Aspin). **Chez les enfants non-chanteurs, en voix parlée**, nous notons des différences significatives entre filles et garçons pour plusieurs paramètres : pour le débit d'air au début du signal (DAB D) $p=0.028$; pour le débit d'air sur la partie stable du signal (DAB S) $p=0.024$. pour les mesures d'écart-type du débit d'air sur la partie stable du signal (DAB ES) $p=0.005$; pour le fondamental moyen au début de l'émission (Fo D) $p=0.048$; pour l'intensité moyenne au début de l'émission (Int D) $p=0.009$; pour l'intensité moyenne sur la partie stable du signal (Int S) $p=0.005$; pour le shimmer de la partie stable du signal (Shim S) $p=0.040$ et pour le shimmer standard exprimé en % au début du signal (Shim std D) $p=0.010$. Les mesures récoltées pour les paramètres précités sont supérieures chez les filles, sauf en ce qui concerne la mesure du fondamental moyen en début d'émission (Fo D). Pour les autres paramètres, nous ne notons pas de différences significatives entre les sexes ($p= N.S.$). Pour la même population, mais en voix chantée, nous observons des différences significatives pour le DAB S $p=0.035$; pour les mesures du coefficient du fondamental sur le début de l'émission (Fcoef D) $p=0.012$; pour Int D $p<0.001$ et pour l'Int S $p<0.001$. Ces mesures sont aussi supérieures chez les filles. Les autres paramètres n'apportent pas de différences significatives inter- sexes ($p = N.S.$).

Pour les enfants chanteurs en voix parlée, il n'y a pas de différences significatives entre les sexes ($p=N.S.$). **En voix chantée**, nous retenons deux résultats significatifs. Les mesures de coefficient d'intensité pour le début de l'émission (Int Coef D) ($p=0.044$) et les mesures de l'écart-type de l'intensité au début de l'émission (Int Écart D) ($p=0.028$) sont supérieures chez les filles. Les paramètres restants ne présentent pas de caractéristiques notables ($p = N.S.$).

Nous avons réalisé une ANOVA afin d'étudier les paramètres suivants (Fo moyen, Fo écart, Fo Coefficient, Intensité moyenne, Intensité écart, Intensité coefficient, Débit d'air buccal, Débit d'air buccal écart, Débit d'air buccal coefficient,

Jitter standard, Jitter pourcent, Shimmer standard, Shimmer %, Shimmer débit et Fuite Glottique) en fonction de la tâche (Parlé vs Chanté), du moment de l'émission (Début vs Stable) et du type d'émission de l'enfant (Chanteur vs Non-chanteur). Cette analyse a été réalisée les deux sexes confondus au vu de l'importance relative du poids du sexe dans le précédent test d'Aspin-Welch (au test de Welch).

Il y a un **effet de la tâche sur la mesure du fondamental** Fo, $p < 0.0001$. Les valeurs de Fo en voix parlée sont inférieures au Fo en voix chantée.

Pour le Fondamental écart (Fo Écart), nous notons **un effet du moment de l'émission** ($p < 0.0001$) et du type d'émission ($p = 0.0003$). Les valeurs de Fo Écart sont supérieures au début de l'émission. Les valeurs du F0 Écart du chanteur sont inférieures à celles du non-chanteur.

Le *Fondamental coefficient* (Fo Coef) durant le début de l'émission est plus élevé que durant sa phase stable ($p < 0.0001$). De plus, le Fo Coef du chanteur est plus bas que celui du non-chanteur, $p = 0.001$. En début d'émission seulement, les valeurs de FC en voix parlée sont supérieures à celles en voix chantée ($p = 0.014$).

Tableau 65 : Valeurs pour les paramètres relatifs à la fréquence fondamentale (Fo).

Fo	Non-chanteurs (N=103)				Chanteurs (N=86)			
Tâches	[a] parlé		[a] chanté		[a] parlé		[a] chanté	
Mom.	Début	Stable	Début	Stable	Début	Stable	Début	Stable
Moy	279.62	278.65	327.59	326.92	271.13	270.08	318.41	317.37
DS	43.12	44.12	96.10	94.39	28.05	30.15	80.92	80.88
Mini	199	203	184	185	213	209	226	224
Maxi	466	465	690	299	379	398	768	762
Fo E.								
Moy	19.06	5.08	19.33	5.49	15.17	2.78	16.64	3.45
DS	11.21	10.77	13.6	5.46	6.89	1.43	8.14	2.43
Mini	3	1	3	0	3	1	3	1
Maxi	59	106	83	35	35	11	42	16
Fo Coef								
Moy	6.917	1.89	5.94	1.79	5.669	1.049	5.224	1.07
DS	4.34	4.48	4.09	2.18	2.662	0.464	2.3	0.653
Mini	1.4	0.5	-3.8	0	1	0.5	1	0.4
Maxi	23.5	44.5	31.3	15.2	14.1	3.4	11.5	5.4

Légende : N : nombre de sujets ; Mom. : moment de l'émission ; Moy. : moyenne ; DS : déviation standard ; Mini : minimum ; Maxi : maximum ; Fo E : fondamental écart ; Fo Coef : fondamental coefficient.

Nous observons un **effet de l'intensité** (Int) ($p < 0.0001$). Les valeurs d'Int des chanteurs sont plus faibles que celles des non-chanteurs.

Les valeurs d'Intensité écart (Int Écart) sont supérieures en voix chantée ($p = 0.003$) et en début d'émission ($p < 0.001$). Elles sont aussi plus élevées pour les chanteurs que pour les non-chanteurs, en début d'émission ($p < 0.001$).

Nous observons les trois mêmes résultats pour le paramètre d'intensité coefficient (Int Coef) (respectivement $p = 0.025$, $p < 0.001$, $p < 0.001$).

Tableau 66 : Valeurs pour les paramètres relatifs à l'intensité (Int).

Int	Non-chanteurs (N=103)				Chanteurs (N=86)			
Tâches	[a] parlé		[a] chanté		[a] parlé		[a] chanté	
Mom.	Début	Stable	Début	Stable	Début	Stable	Début	Stable
Moy	77.55	77.12	78.58	78.17	72.23	71.82	72.33	72.25
DS	6.44	6.66	8.05	8.26	5.15	4.98	5.55	6.24
Mini	56.78	58.54	50.83	55.53	54.42	59.86	59.43	59.95
Maxi	90.73	90.94	94	95.5	87.52	87.67	84.02	85.58
Int E.								
Moy	4.13	0.77	4.58	1.21	5.02	0.76	5.2	0.87
DS	1.93	0.44	2.57	1.9	1.44	0.61	1.55	0.58
Mini	0.76	0.25	0.76	0.25	1.88	0.12	2.08	0.2
Maxi	8.51	3.04	15	16.27	11.76	3.77	8.52	4.13
Int Coef								
Moy	5.33	1.02	5.67	1.63	6.96	1.07	7.14	1.23
DS	2.25	0.64	3	2.8	2.18	0.87	1.9	0.87
Mini	1.1	0.3	1	0.4	2.8	0.2	3.3	0.3
Maxi	9.6	4.3	21	24.1	21.6	5.8	11.9	5.9

Légende : N : nombre de sujets ; Mom. : moment de l'émission ; Moy. : moyenne ; DS : déviation standard ; Mini : minimum ; Maxi : maximum ; Int E : intensité écart ; Int coef : intensité coefficient.

Ensuite, nous notons un **débit d'air buccal (DAB)** plus important en début d'émission ($p < 0.0001$) chez les chanteurs ($p < 0.0001$). Chez un non-chanteur, les valeurs de DAB seront plus faibles en voix parlée qu'en voix chantée ($p = 0.024$).

Pour le Débit d'air buccal écart (DAB Écart), il y a un effet principal du moment de l'émission ($p < 0.0001$). Les valeurs du DAB écart sont supérieures en début d'émission. Nous notons une interaction significative des trois variables ($p = 0.04$). Les valeurs de DAB écart sont plus faibles en voix parlée qu'en voix chantée pour les non-chanteurs au début de l'émission ($p = 0.003$). Elles sont aussi plus faibles chez les

chanteurs que chez les non-chanteurs au début de l'émission et en voix chantée ($p=0.013$).

Ensuite, l'analyse du débit d'air coefficient (DAB Coef) nous indique une interaction significative entre le moment de l'émission et le type d'émission du chanteur ($p=0.005$). Les valeurs de DAB Coef sont plus élevées chez le chanteur au début de l'émission qu'en phase stable d'émission ($p<0.0001$). En outre, le chanteur a des valeurs de DAB Coef plus importantes au début de l'émission que le non-chanteur ($p<0.0001$).

Tableau 67 : Valeurs pour les paramètres relatifs au débit d'air buccal (DAB).

DAB	Non-chanteurs (N=103)				Chanteurs (N=86)			
Tâches	[a] parlé		[a] chanté		[a] parlé		[a] chanté	
Mom.	Début	Stable	Début	Stable	Début	Stable	Début	Stable
Moy	0.008	0.002	0.016	0.011	0.066	0.056	0.062	0.54
DS	0.085	0.085	0.089	0.09	0.041	0.038	0.042	0.04
Mini	-0.18	-0.18	-0.17	-0.17	0	-0.04	0	-0.01
Maxi	0.15	0.15	0.3	0.21	0.23	0.2	0.18	0.18
DAB E								
Moy	0.014	0.006	0.019	0.007	0.017	0.005	0.014	0.004
DS	0.009	0.006	0.022	0.007	0.012	0.007	0.011	0.006
Mini	0	0	0	0	0	0	0	0
Maxi	0.04	0.02	0.2	0.04	0.07	0.05	0.05	0.02
DAB C								
Moy	10.77	8.05	12.04	8.09	26.55	10.22	24.75	13.54
DS	35.48	18.58	62.45	16.7	22.37	13.99	47.69	14.88
Mini	-242.2	-16.3	-456	-53.7	-89.1	-71.4	-328	-14.3
Maxi	107.2	164.8	306	72.7	148.9	63.2	256	117

Légende : N : nombre de sujets ; Mom. : moment de l'émission ; Moy. : moyenne ; DS : déviation standard ; Mini : minimum ; Maxi : maximum ; DAB E : débit d'air buccal écart ; DAB C : débit d'air buccal coefficient.

Le **paramètre Jitter** (Jitt) est plus élevé au début de l'émission que pendant la phase stable ($p<0.0001$). Les valeurs de Jitt sont également plus petites chez le chanteur que chez le non-chanteur au début de l'émission ($p<0.0001$).

Les mêmes résultats sont observés pour le paramètre Jitter% (respectivement $p<0.0001$ et $p<0.0001$).

Tableau 68 : Valeurs pour les paramètres relatifs à la variabilité de la fréquence fondamentale : le Jitter (Jitt).

Jitter	Non-chanteurs (N=103)				Chanteurs (N=86)			
Tâches	[a] parlé		[a] chanté		[a] parlé		[a] chanté	
Mom.	Début	Stable	Début	Stable	Début	Stable	Début	Stable
Moy	5.09	1.85	5.09	2.06	3.34	1.19	3.39	1.57
DS	4.57	2.18	6.57	2.48	2.06	0.38	1.63	1.47
Mini	1.7	0.5	1	0	1	0.7	1.1	0.7
Maxi	29.9	19.9	43.4	19.8	16.7	2.4	9.6	14.3
Jitter %								
Moy	1.87	0.72	1.64	0.69	1.24	0.44	1.11	0.48
DS	1.77	1.01	2.2	1.02	0.76	0.14	0.55	0.25
Mini	0.4	0.2	0.2	0	0.4	0.2	0.3	0.2
Maxi	11.8	7.2	16.4	8.6	6.1	1	3.1	2.4

Légende : N : nombre de sujets ; Mom. : moment de l'émission ; Moy. : moyenne ; DS : déviation standard ; Mini : minimum ; Maxi : maximum ; jitt % : jitter exprimé en pour-cent.

Nous observons **un effet du moment de l'émission pour le shimmer** (Shim). Ce dernier est plus élevé en début d'émission qu'en phase stable ($p < 0.0001$). Nous retenons que les valeurs de shimmer sont plus importantes en voix parlée qu'en voix chantée dans la population des non-chanteurs en début d'émission ($p < 0.0001$) et en phase stable ($p = 0.025$). Enfin, les valeurs de shim du chanteur sont plus faibles que celles du non-chanteur en début d'émission en voix parlée ($p < 0.0001$), en phase stable parlée ($p < 0.0001$) et chantée ($p < 0.0001$).

Pour le *Shimmer dB* (Shim dB) nous avons aussi un effet du moment de l'émission ($p < 0.0001$). Le début de l'émission récolte des mesures plus importantes que la fin. Le Shim dB est plus élevé en voix parlée que chantée en début d'émission chez le chanteur ($p < 0.0001$) et durant toute la durée de l'émission chez le non-chanteur (respectivement $p < 0.0001$ et $p = 0.025$). Les valeurs de Shim dB restent plus faibles chez le chanteur que chez le non-chanteur en voix parlée ($p < 0.0001$) et chantée ($p = 0.002$) durant la partie stable de l'émission.

Le *shimmer %* est plus élevé en début d'émission qu'en phase stable ($p < 0.0001$). Il y a une interaction significative entre la tâche et le type d'émission des enfants ($p < 0.039$). Les valeurs de shim% sont plus élevées en voix parlée qu'en voix chantée chez les chanteurs ($p = 0.014$). De plus, nous observons des valeurs de shim%

inférieures chez le chanteur par rapport au non-chanteur en voix parlée ($p=0.033$) et en voix chantée ($p<0.0001$).

Quant au *shimmer débit* %, il ne présente aucune particularité discutable.

Tableau 69 : Valeurs pour les paramètres relatifs aux mesures de variabilité de l'intensité : le Shimmer (shim).

Shim	Non-chanteurs (N=103)				Chanteurs (N=86)			
Tâches	[a] parlé		[a] chanté		[a] parlé		[a] chanté	
Mom.	Début	Stable	Début	Stable	Début	Stable	Début	Stable
Moy	0.17	0.08	0.15	0.07	0.14	0.03	0.14	0.03
DS	0.05	0.05	0.05	0.06	0.05	0.05	0.06	0.05
Mini	0.1	0	0.1	0	0.1	0	0.1	0
Maxi	0.4	0.2	0.2	0.2	0.3	0.2	0.3	0.1
Shim %								
Moy	0.22	0.1	0.18	0.09	0.21	0.06	0.18	0.06
	0.07	0.06	0.07	0.07	0.07	0.07	0.07	0.06
DS								
Mini	0.1	0	0.1	0	0.1	0	0.1	0
Maxi	0.5	0.3	0.5	0.3	0.4	0.2	0.4	0.2
Shim DB								
Moy	0.77	0.52	0.8	0.52	0.71	0.45	0.62	0.41
DS	0.28	0.13	0.43	0.2	0.22	0.13	0.29	0.15
Mini	0.09	0.28	0.26	0	0.34	0.25	0.17	0.15
Maxi	1.77	0.92	3	1.17	1.58	0.93	1.61	0.97

Légende : N : nombre de sujets ; Mom. : moment de l'émission ; Moy. : moyenne ; DS : déviation standard ; Mini : minimum ; Maxi : maximum ; shim% : shimmer exprimé en pour-cent ; shim DB : shimmer débit.

Enfin, nous observons **un effet du moment de l'émission pour la Fuite Glottique** (FG), plus faible au début ($p<0.0001$). En outre, chez les non-chanteurs, les valeurs de FG sont moindres en voix parlée qu'en voix chantée tant au début de l'émission ($p=0.021$) qu'en phase stable ($p<0.0001$). Chez les chanteurs, les valeurs de FG sont inférieures à celles des non-chanteurs en voix parlée ($p<0.0001$) et chantée ($p=0.0001$), seulement en phase stable d'émission.

Tableau 70 : Valeurs pour les paramètres relatifs à la fuite glottique (FG).

FG	Non-chanteurs (N=103)				Chanteurs (N=86)			
Tâches	[a] parlé		[a] chanté		[a] parlé		[a] chanté	
Mom.	Début	Stable	Début	Stable	Début	Stable	Début	Stable
Moy	4.66	7.39	5.54	8.85	5.03	11.96	4.76	11.23
DS	1.83	3.5	3.27	5.18	1.7	6.61	2.1	6.28
Mini	1.94	0.97	1.85	0	1.53	2.18	1.83	2.87
Maxi	11.99	21.58	24.66	33.87	10.04	30.28	11.54	32.57

Légende : mom : moment de l'émission ; Moy : moyenne ; DS : déviation standard ; Mini : minimum ; Maxi : maximum.

5.4.4. Discussion et conclusion.

La fiabilité test/re-test atteste de notre fiabilité dans l'exploitation du matériel, la concordance étant bonne pour tous les paramètres. Au vu des données statistiques, les valeurs obtenues nous permettent d'aboutir à une standardisation des paramètres acoustiques et aérodynamiques. La taille de l'échantillon s'avère suffisante tant pour le groupe d'enfants non-chanteurs que pour le groupe d'enfants chanteurs.

Quatre variables chacune à deux niveaux sont introduites dans l'étude,

- le sujet : enfants non-chanteurs versus enfants chanteurs, le genre
- le type d'émission : voix parlée versus voix chantée
- le moment de l'émission : début du signal versus partie stable du signal.

- **Concernant le genre :**

Les filles sans entraînement vocal présentent des valeurs supérieures par rapport aux garçons pour les mesures de débit d'air et d'intensité tant en voix parlée qu'en voix chantée.

Les choristes présentent des valeurs supérieures mais uniquement en voix chantée pour les mesures de variabilité d'intensité (écart et coef) et seulement pour le début de l'émission.

Contrairement aux idées reçues, les petites filles consomment plus d'air que les petits garçons. Elles ont un comportement plus tonique et pourtant elles ont moins de lésions de malmenage vocal.

Elles ont un meilleur contrôle lorsqu'elles sont entraînées au chant excepté pour l'attaque où il y a un temps d'adaptation de l'intensité. Elles sont soumises à la technique de l'examen, il s'agit donc de gestes moins naturels que dans la vie courante.

Culture, timidité, comportement du groupe, spontanéité doivent être pris en compte.

- **Concernant la tâche :**

La tâche parlée entraîne des variations sur les mesures de débit d'air et toutes les mesures de shimmer. L'effet de cette tâche est fonction des sujets et des moments d'émission.

La tâche chantée a un impact sur le débit d'air buccal, le shimmer et l'intensité écart. Les valeurs sont supérieures, la voix chantée entraînant une consommation d'air plus importante, la voix chantée consomme plus et est spontanément plus intense, mais elle est contrôlée et beaucoup plus stable.

- **Concernant les sujets :**

Les non-chanteurs présentent des valeurs plus élevées pour les mesures de variabilité de la fréquence (Fo écart, Fo coef, jitter et jitter % uniquement pour le début), de variabilité de l'intensité (shim et shim%), d'intensité et de fuite glottique uniquement pour la partie stable. [on peut penser à un moins bon contrôle] .Par contre, ils présentent des valeurs inférieures au débit d'air buccal en voix chantée.

Les chanteurs quant à eux ont des valeurs supérieures aux mesures d'intensité et au débit d'air buccal mais uniquement en voix chantée. Ils montrent également des valeurs de variabilité d'intensité (int écart et coef) supérieures pour le début de l'émission. Le fait que les choristes donnent quantitativement plus d'intensité vocale pourrait être dû à une première étape dans leur conscientisation de la tâche chantée et de ses exigences. Cependant leur contrôle de la coordination pneumophonique ne s'affirmera que progressivement avec l'âge, la maturité et la motivation.

- **Concernant le moment de l'émission :**

Le début de l'émission ou l'attaque a un impact sur les mesures de variabilité de la fréquence (Fo écart, jitter et jitter %), de variabilité de l'intensité (shim % et shim dB), et de débit d'air écart. Les valeurs sont supérieures par rapport au moment stable de l'émission pour les deux groupes de sujets. La comparaison des paramètres acoustiques et aérodynamiques des deux moments de l'émission nous montrent des valeurs supérieures pour le début de l'émission. Au moment de l'attaque du son, l'enfant "lâche" sa voix brusquement. L'équilibre et la précision de la mise en vibration (voice onset) sont le résultat d'un travail neuro-musculaire complexe. Nous supposons que son bagage anatomophysiologique ne lui permet pas encore la complète maîtrise de son geste. On assiste dès lors à de l'irrégularité à l'attaque. La fuite glottique, quant à elle est moins importante, révélant peut-être plus de contraction à l'attaque. Rantala et al (1997)^[245] analysent également l'impact du moment de l'émission sur les mesures acoustiques et rapportent un effet sur le shimmer et la hauteur tonale. Ils précisent que les différences entre les deux moments sont minimales, mais néanmoins statistiquement significatives. Ils observent des différences dans la variation de la fréquence fondamentale, variation plus importante à la fin de l'émission. Ils interprètent ces différences comme dues à un meilleur contrôle de la stabilité en début d'émission. Ses valeurs de jitter et de shimmer sont légèrement plus élevées en fin d'émission de la voyelle, mais ces différences ne sont pas significatives. Ce changement dit-il, peut provenir d'une diminution probable de la poussée trachéale laquelle à son tour entraîne un changement dans l'ajustement laryngé.

Nous ne retrouvons pas ces résultats dans notre étude. Les valeurs y sont plus importantes sur le début de l'émission, ce qui rend contestable l'intérêt de cette comparaison avec Rantala. Cependant, nous n'avons pas pris en considération les valeurs de fin d'émission, mais les valeurs sur émissions stables. Nous observons plutôt que l'attaque déstabilise les mesures de variabilité, ce qui corroborait les conseils de Schoentgen (1989)^[277] sur les perturbations qu'entraîne l'attaque sur le jitter et le shimmer. Les enfants éprouveraient des difficultés à maîtriser la pression sous glottique régularisant la fréquence et l'intensité. Les valeurs de débit d'air sont supérieures pour les chanteurs uniquement à l'attaque. En chorale, chanter ensemble correspond à démarrer ensemble. Les choristes sont conditionnés à réaliser une prise

d'air en fonction de la production vocale qu'ils doivent réaliser. Emmagasiner juste la quantité adéquate est le fruit d'un long apprentissage.

Les valeurs de fuite glottique sont supérieures sur la partie stable de l'émission.

5.5. ÉTUDE 6 : APPLICATION DE LA STATION E.V.A.1 POUR L'ÉVALUATION OBJECTIVE DES DYSPHONIES APRÈS PARALYSIE D'UNE CORDE VOCALE.

5.5.1. Introduction.

L'objectif de cette étude est de contribuer à déterminer les mesures acoustiques et aérodynamiques fournies par l'appareillage E.V.A 1^[293] utilisable pour les voix en les comparant aux mesures perceptives G.I.R.BA.S.^[122] fournies par un jury d'écoute, à partir d'un groupe de patients présentant une même pathologie, à savoir la paralysie unilatérale de la corde vocale.

5.5.2. Matériel et méthode.

Nous avons enregistré les voix de 40 sujets répartis en deux groupes. Le premier groupe était composé de 28 sujets, 10 hommes et 18 femmes, atteints d'une paralysie laryngée latérale. Le deuxième groupe comportait 12 sujets sains, 7 hommes et 5 femmes. Il constituait le groupe témoin. Le jury d'écoute était composé de 8 auditeurs experts, 2 oto-rhino-laryngologues, 1 phoniatre, et 5 orthophonistes. Les enregistrements vocaux ont été réalisés à l'aide d'un enregistreur numérique de type D.A.T. (SONY WALKMAN TCD-D7 (SONY corporation, Tokyo, Japon) et d'un microphone sur pied placé à 30 cm de la bouche du sujet (SONY ECM-727P stéréophonique et unidirectionnel). Les auditeurs ont évalué la qualité vocale des sujets sur un extrait des "*Lettres de mon moulin*" d'A. Daudet : "*La chèvre de Monsieur Seguin*". Ils inscrivaient un signe clair sur une échelle analogique visuelle de 10 cm. Les mentions "très bon" et "très mauvais" étaient portées de part et d'autre de l'échelle. Chaque graduation était fournie avec une explication claire des extrêmes. Les mesures recueillies sur cette échelle ont été retenues avec une précision d'un centimètre. Les échantillons vocaux étaient évalués selon les six critères de l'échelle

G.I.R.B.A.S (de Krom et al (1994)^[54], (1995)^[55], Dejonckere (1995)^[62], Murry et al (1977)^[209]). Les auditeurs répondaient également à la question "Ce sujet est-il atteint d'une paralysie laryngée ? " par oui, non, je ne sais pas. Chaque stimulus a été présenté trois fois ou plus dans un endroit calme. Les sessions d'écoute en fonction des disponibilités du jury ont parfois eu lieu en petit groupe de 2. voire 3 à 5 auditeurs.

L'enregistrement des signaux acoustiques a été réalisé par un microphone coaxial AKG 409 contenu dans l'embouchure et placé à 5 cm du sujet. Les sujets devaient produire les syllabes "papa" et tenir le dernier [a] quelques secondes. Pour les stimuli à évaluer par le jury d'écoute, nous avons choisi la lecture d'un texte. Pour l'analyse objective avec le logiciel E.V.A. ^[293], nous avons opté pour la voyelle tenue. L'examineur retenait l'essai qui lui semblait le plus proche de la voix spontanée du sujet. Les différents paramètres acoustiques et aérodynamiques, (le fondamental (Fo), le coefficient du fondamental (Fo coef), l'intensité (Int), le coefficient de l'intensité (int coef), le débit d'air buccal (DAB), la fuite glottique, le jitter (Jitt), et le shimmer (Shim)) ont été calculés par le logiciel E.V.A sur une fenêtre de deux secondes sélectionnées dans la partie stable de la voyelle, 300 millisecondes après l'attaque (Giovanni (1995)^[94]).

Pour chacun des six paramètres subjectifs (G.I.R.B.A.S.), nous avons réalisé une analyse de variance (Anova) des mesures répétées par une approche multi-variée (modèle aléatoire, test T² de Hotelling) selon trois critères : la session d'écoute, les juges (n=8), et les sujets (pathologiques/témoins). Nous avons regardé s'il y avait l'effet d'un critère (inter-juges, intra-juges, sujets (pathologiques/témoins)), un effet des critères 2 à 2 (inter-juges x intra-juges, sujets x inter-juges ; sujets x intra-juges) , ou un effet des trois critères (sujets x inter-juges x intra-juges) Pour la fiabilité des mesures perceptives recueillies sur l'échelle G.I.R.B.A.S., nous avons réalisé une analyse de variance (ANOVA). Pour la validité des réponses, nous avons calculé l'index surface donné par la courbe Roc. Pour la corrélation des paramètres perceptifs entre eux , nous avons utilisé le r de Pearson et pour celle qui apparaîtrait entre les paramètres objectifs et les mesures perceptives, nous avons appliqué le r de Spearman.

5.5.3. Résultats.

Le tableau 71 montre la concordance intra-juges (3 lectures) pour les 8 juges quant au diagnostic de présence éventuelle d'une pathologie laryngée. Le taux médian de concordance est globalement de 75%, soit 75% pour les pathologies laryngées et 83% pour le groupe témoin (NS).

Tableau 71 : Concordances intra-juges aux 3 lectures.

Jury	Pathologies (n=28)		Témoins (n=12)	
	%	N	%	N
1	88	21	75	6
2	82	22	79	11
3	82	26	93	7
4	78	21	75	10
5	65	15	54	11
6	82	23	82	10
7	65	20	71	6
8	72	19	68	10
Médiane	75		75	

Le tableau 72 montre le nombre de concordances entre les juges, en utilisant pour chacun d'eux le résultat le plus constant des 3 lectures. Cette concordance est parfaite dans 58 % des cas et ne diffère pas significativement entre les pathologies laryngées et les témoins.

Le diagnostic de pathologie laryngée est posé par chacun des 8 juges avec une sensibilité variant entre 54 et 96 % (médian 84 %) et une spécificité variant entre 67 et 100 % (médian 92 %).

Si l'on utilise le nombre total de réponses en faveur de la pathologie laryngée pour l'ensemble du jury (0 à 8), la valeur diagnostique mesurée par la surface sous la courbe ROC est de 0.984 +/- 0.016.

Tableau 72 : Nombre de concordances entre les juges.

Nbre maximal de concordances	Tous	Pathologies	Témoins
8	23 (58%)	15	8
7	8 (20%)	6	2
6	4 (10%)	4	-
5	4 (10%)	2	2
4	1 (2%)	1	-

Le tableau 73 détaille le résultat de l'analyse de variance appliquée aux 6 paramètres de l'échelle G.I.R.B.A.S.. Il n'existe pas de différences statistiquement significatives entre les lectures d'un même juge. Pour la faiblesse (A), les juges diffèrent entre eux ($p = 0.025$). Pour les autres paramètres, ils diffèrent encore entre eux dans l'appréciation des voix pathologiques ($p < 0.001$) mais pas pour celles des témoins. Par ailleurs, les mesures des 6 paramètres sont plus élevées pour les voix pathologiques que pour les témoins ($p < 0.001$).

Tableau 73 : Analyse de variance des 6 paramètres de l'échelle G.I.R.B.A.S..

Paramètres	INTERACTIONS							
	P*, inter, intra	Inter, intra	P*, inter	P*, intra	Inter-juges		Intra-juges	
					P*	T**	P*	T**
Grade	N.S.	N.S.	P=0.008	N.S.	P<0.001	N.S.	N.S.	P<0.001
Instabilité	N.S.	P=0.033	P=0.001	N.S.	P<0.001	N.S.	N.S.	P<0.001
Raucité	N.S.	P=0.001	P=0.001	N.S.	P<0.001	N.S.	N.S.	P<0.001
Souffle	N.S.	N.S.	P=0.005	N.S.	P<0.001	N.S.	N.S.	P<0.001
Faiblesse	N.S.	N.S.	N.S.	N.S.	P=0.025		N.S.	P<0.001
Forçage	P=0.008	P=0.008	P<0.001	N.S.	P<0.001	N.S.	P<0.001	P<0.001

Légende : P*= pathologiques, T**= témoins.

Le tableau 74 montre la bonne corrélation qui existe entre les 6 paramètres subjectifs, tant pour les voix pathologiques ($p < 0.001$) que pour les témoins ($p < 0.01$) sauf en ce qui concerne la corrélation entre les mesures de raucité (R) et de forçage (S), pour les témoins ($p = 0.064$).

Tableau 74 : Corrélations entre les paramètres de l'échelle G.I.R.B.A.S. (Spearman).

	Témoins (n=12) Coefficient de corrélation					
Paramètres	Grade	Instabilité	Raucité	Souffle	Faiblesse	Forçage
Grade (G)	1.000					
Instabilité (I)	.977	1.000				
Raucité (R)	.885	.888	1.000			
Souffle (B)	.956	.919	.909	1.000		
Faiblesse (A)	.897	.908	.774	.862	1.000	
Forçage (S)	.712*	.722*	.550**	.707*	.882	1.000
	Pathologiques (n=28) Coefficient de corrélation					
Paramètres	Grade	Instabilité	Raucité	Souffle	Faiblesse	Forçage
Grade (G)	1.000					
Instabilité (I)	.904	1.000				
Raucité (R)	.750	.900	1.000			
Souffle (B)	.961	.837	.661	1.000		
Faiblesse (A)	.963	.845	.644	.980	1.000	
Forçage (S)	.969	.951.	.826	.916	.928	1.000

Légende : P<0.001 excepté pour * et ** ; *p≤0.01 ; **p=0.064

Le tableau 75 détaille les corrélations entre les paramètres subjectifs et les paramètres objectifs. Les mesures les plus corrélées aux paramètres subjectifs sont le coefficient de variabilité fondamentale (Fo coef, p< 0.001), le jitter (p< 0.001) et le coefficient de variabilité de l'intensité (Int coef, p≤ 0.002). L'intensité moyenne, le DAB moyen, le coefficient de variabilité du DAB sont corrélés avec certains paramètres subjectifs, mais il n'y a pas de corrélation statistiquement significative des paramètres subjectifs avec le fondamental moyen, le shimmer dB et le shimmer DAB.

Tableau 75 : Corrélations entre les paramètres subjectifs et objectifs (Spearman's rho).

Mesures	Grade	Instabilité	Raucité	Souffle	Faiblesse	Forçage
Fo moyen	.140	.237	.198	.142	.173	N.S.
Fo écart	.671 / .000	.662 / .000	.704 / .000	.633 / .000	.601 / .000	.687 / .000
Fo coef	-.401 / .011	-.308	-.236	-.389 / 0.14	-.404 / .011	-.349 / .029
Int moy.	.530 / .011	.537 / .000	.525 / .001	.485 / .002	.480 / .002	.480 / .002
DAB	.351 / .028	.290	.285	.363 / .023	.378 / .018	.353 / .027
DAB coef	.346 / .031	.303	.375 / .019	.303	.295	.312
Jitter	.615 / .000	.569 / .000	.622 / .000	.592 / .000	.561 / .000	.590 / .000/
Shim dB	.246	.161	.104	.224	.257	.213
Shim DAB	.196	.135	.137	.194	.182	.201

Légende : .000 : p<0.001

5.5.4. Discussion et Conclusion.

- **Les mesures perceptives**

Le choix du jury d'écoute, de l'échelle d'évaluation, des extraits vocaux, de la population étudiée sont quatre facteurs, sources éventuelles de variabilité pour les mesures perceptives.

En ce qui concerne les critères pour le choix du jury, certains auteurs (Murry et al (1977)^[209] ; Kempster et al (1991)^[161] ; Kreiman et al (1990)^[169]) n'ont observé aucune différence entre les auditeurs naïfs et les experts. L'explication serait que le protocole montre une sensibilité insuffisante. Kreiman et al (1990)^[169] ont rapporté dans leur étude que pour les jugements des voix pathologiques les auditeurs développaient des approches idiosyncratiques dans leur pratique clinique. Ils stockeraient en mémoire à long terme des exemplaires typiques pour des qualités perceptives variées. Ils ajoutaient que les auditeurs experts ont besoin de plusieurs années pour développer une série de critères stables pour évaluer les voix. Les auditeurs naïfs, par contre, manqueraient de modèles internes spécifiques. Ils font référence à leur propre modèle de voix. Pour éviter l'écueil d'un jury naïf inadéquat, notre choix s'oriente en faveur des auditeurs experts.

Pour le choix de l'échelle d'évaluation, Kreiman et al (1993)^[171] dans leur étude ont fait une revue de celles qu'on trouve dans la littérature. L'échelle graduée, les auditeurs assignent un nombre de 1 à n (nombre de graduation sur l'échelle) aux échantillons de voix. L'échelle la plus souvent employée pour l'évaluation vocale a un nombre de points équidistants de 5 à 10. L'ordre de présentation des stimuli influence le choix de la valeur sur l'échelle. Par exemple, la sensation de faiblesse apparaît comme étant plus forte à la deuxième représentation. Gerratt et al (1993)^[91] ont confirmé cette observation. Nous avons opté pour une échelle analogique visuelle.

Lors de leur évaluation perceptive, les dysphonies sont classées et décrites en termes d'impressions auditives (voix soufflées, enrouées, sèches, humides, rauques...) Hirano (1989)^[125] rapporte que la définition de ces termes a toujours donné lieu à controverse entre spécialistes. D'où la nécessité de standardiser l'évaluation psycho-acoustique des voix pathologiques et la terminologie qui s'y rapporte. Des auteurs (Geratt et al^[91] ; Giovanni et al^[92-98], Hammarberg et al^[108] ; Hirano^[125]) se sont attelés à établir des caractéristiques vocales perceptives. Le

Comité pour les tests de la fonction phonatoire de la Société Japonaise des Logopèdes et Phoniâtres a proposé l'échelle G.R.B.A.S. pour l'évaluation des dysphonies. de Krom ^[54. 55] a utilisé également le facteur I se rapportant à l'instabilité du signal. Le groupe G.R.E.L. (Groupe de Recherche Européen sur le Larynx) a proposé également d'ajouter ce facteur à l'échelle^[62]. Bien que ces 6 paramètres perceptifs (Grade, Raucité, Souffle, sensation de Faiblesse, Forçage, Instabilité) ne soient pas suffisants pour rendre compte de toutes les dimensions du trouble vocal, beaucoup d'auditeurs ont considéré ces derniers comme prédominants et constants (de Krom et al^[55]). Pour les paramètres de l'échelle G.I.R.B.A.S., les auditeurs ont pu manquer d'éléments quant à la définition des paramètres à évaluer. Une définition plus complète de chacun de ces critères uniformiserait sans doute l'évaluation.

Nous avons choisi un texte à lire pour l'évaluation par jury d'écoute, et l'émission de la voyelle [a] pour l'analyse acoustique.

Les avantages de la voyelle tenue résident dans le fait que sa production est plus facilement contrôlée et standardisée que la parole. L'auditeur se concentre mieux sur le signal. De plus, cette échelle répond mieux aux exigences de l'analyse acoustique pour l'obtention de paramètres fiables, de par la régularité et la stabilité du signal. Cependant, la parole est plus représentative que la voyelle. Les attaques, les terminaisons vocaliques, les pauses, les segments de transition peuvent être analysés^[108]. Les modèles musculaires laryngés sont plus complexes et dynamiques dans la parole que pour la voyelle. Ils sont donc susceptibles de mettre en évidence les aspects déviants de la voix, particulièrement l'immobilité cordale rendant les évaluations perceptives plus fiables. Grâce aux facteurs linguistiques segmentaux et suprasegmentaux inhérents au langage oral, on peut obtenir une description plus détaillée des aspects pathologiques.

Cependant les éléments de la parole continue seraient plus confondants qu'informatifs dans l'évaluation subjective de la qualité vocale. L'auditeur serait encombré de données inutiles telles que l'accent, l'aspect sémantique du message, les patterns articulatoires. de Krom et al^[55] ont démontré la fiabilité des mesures perceptives concernant le Grade, la Raucité, le Souffle. Elles étaient meilleures pour les fragments vocaliques comprenant l'attaque. Il serait donc intéressant d'utiliser un extrait vocalique incluant l'attaque.

Le type de pathologie choisi est la paralysie unilatérale de la corde vocale. Lors d'une paralysie laryngée unilatérale, la corde vocale peut être immobilisée en différentes positions par rapport à la ligne médiane. La dysphonie sera d'autant plus sévère que l'écartement des cordes vocales par rapport à la ligne médiane est important. Les relations entre lésions neurologiques et position des cordes vocales ont fait l'objet de plusieurs études^[54, 122]. Ces relations ne sont pas absolues. La gravité de la dysphonie peut varier d'un enrouement léger à une aphonie quasi totale. L'atrophie du muscle vocal, la mise en place de mécanismes compensatoires et la régénérescence du nerf atteint sont autant de facteurs qui modifient la dysphonie. Cependant comparée avec des pathologies telles que nodules, polypes, sulcus, la paralysie unilatérale présente un trouble de fonctionnement des cordes vocales qui n'altère pas la morphologie de celles-ci.

- **Fiabilité des auditeurs.**

Les accords intra et inter-juges sur la question de savoir, si tel sujet est atteint d'une paralysie laryngée ou non, sont bons. L'accord intra-juges montre que les juges sont constants dans leur évaluation des deux groupes avec une concordance légèrement meilleure pour le groupe témoin. Les auditeurs offrent une bonne concordance entre eux pour la discrimination des sujets (pathologiques/témoins). Ils déterminent relativement bien le statut réel des sujets. Nous pensons qu'un auditeur évalue une qualité perceptive de manière plus cohérente lorsque le critère à évaluer est mieux distinct des autres. Cette observation confirme l'utilité de réaliser cette expérience sur une population de patients présentant la même pathologie vocale, donc des sujets susceptibles de présenter un tableau perceptif comparable. Les résultats sont moins bons selon les types de voix (normales/pathologiques). La variabilité de jugement porte sur les voix dont le degré de dysphonie se situe entre deux extrêmes. Ces constatations rejoignent celles d'autres auteurs comme Kreiman et al (1992)^[170] et Gerratt et al (1993)^[91]. Gerratt et al ont émis l'hypothèse que ce comportement perceptif pouvait être expliqué par un modèle prototypique de la perception de la qualité vocale proposé par Papcun et coll (1989)^[227]. Ce modèle suggère la création par chaque auditeur d'un prototype en cours d'expérience perceptive. Ces prototypes sont donc instables de manière inhérente puisqu'ils sont continuellement en remaniement et peuvent être influencés par des facteurs internes

tels que les trous de mémoire, les problèmes d'attention, etc... Gerratt et al^[91] ont essayé de contrôler cette variabilité prototypique d'un auditeur à l'autre en proposant des stimuli d'ancrage correspondant à un degré de raucité précis allant de 1 à 5 (stimuli synthétiques). Ils ont présenté ces stimuli aux auditeurs avant leur tâche d'évaluation de la raucité de voix synthétiques. Pour ces évaluations, les auditeurs montraient une meilleure fiabilité à propos de la raucité des voix localisées près des "stimuli-prototypes" et une fiabilité moins bonne pour les voix situées entre ceux-ci. Kreiman et al^[171] recensent une série d'études portant sur l'évaluation vocale, d'où il ressort que la fiabilité des auditeurs fluctue grandement d'étude en étude, de même que les protocoles d'évaluation et les méthodes varient d'une étude à l'autre.

Les paramètres perceptifs de souffle et de raucité sont fortement utilisés en clinique. Kreiman et al (1994)^[172] constatent qu'il n'y a aucun consensus sur leur définition. Ce manque de réalité perceptive de certains paramètres serait une des sources de variabilité de l'évaluation. Ainsi, nos résultats montrent que pour le critère de sensation de faiblesse, les juges sont cohérents de manière interne (accord intra-juges), mais qu'ils ne présentent aucune fiabilité inter-juges. Ils ont donc bien chacun leur conception personnelle de ce critère. La notion de sensation de faiblesse peut être comprise comme faible intensité sonore ou encore potentialité de ne pas pouvoir émettre un son de forte intensité. Nous observons un effet des trois critères (Fo coef, jitter et int coef) pour le paramètre de forçage vocal. Les résultats de ce paramètre présentent une différence significative entre les sessions, d'un auditeur à l'autre et selon les types de voix. Ces résultats pourraient signifier que le paramètre de forçage vocal n'est pas un concept clair et fixe chez tous les auditeurs puisque les mesures varient non seulement d'un auditeur à l'autre mais également d'une session d'écoute à l'autre. Dans notre étude, nous remettons en question la validité de ce paramètre dans notre étude. Aussi les résultats qui s'y rapportent ne peuvent être pris en compte dans la discussion des caractéristiques perceptives de la paralysie laryngée et de la comparaison des mesures perceptives avec les mesures objectives.

Pour augmenter la fiabilité lors des évaluations perceptives de la qualité vocale par les auditeurs, il serait nécessaire de définir de manière cohérente les paramètres. De même, il serait pertinent d'avoir recours à une échelle d'évaluation à référents externes.

Les mesures perceptives de grade, de souffle, et de sensation de faiblesse obtiennent les moyennes les plus élevées. Les corrélations obtenues entre ces

mesures indiquent que le grade est fortement corrélé au souffle et à la sensation de faiblesse. Ces informations peuvent signifier que les dimensions perceptives influençant le plus le jugement de la raucité chez les paralysés laryngés sont le souffle et la faiblesse. Dejonckere et al (1986)^[59] rapportent d'ailleurs que l'impression subjective de la sévérité de la dysphonie chez des sujets atteints de diverses pathologies vocales est plutôt liée au caractère soufflé qu'à l'irrégularité de la fréquence fondamentale. Les mesures de raucité et d'instabilité sont bien corrélées entre elles. Les différentes dimensions perceptives s'organiseraient en deux groupes de paramètres liés entre eux : d'une part le grade, le souffle et la faiblesse ; d'autre part la raucité et l'instabilité.

- **L'Analyse Objective.**

Les mesures de variabilité du signal vocal (Fo coef, INT coef, et Jitter) sont bien corrélées avec les critères de l'échelle G.I.R.B.A.. Le forçage n'a pas été repris à cause de la trop grande diversité des jugements et le peu de reproductibilité constaté au cours des trois sessions d'écoute. Par contre les mesures de fondamental moyen, et de shimmer ne trouvent pas de correspondant significatif avec les critères subjectifs.

Les modifications éventuelles de la fréquence fondamentale, du pourcentage de variabilité de l'intensité et du débit d'air buccal ne semblent pas jouer un rôle dans l'appréciation des paralysies unilatérales.

Les dimensions perceptives les plus présentes chez les patients atteints de paralysie laryngée c'est-à-dire le grade, le souffle et la sensation de faiblesse trouvent des correspondants significatifs parmi les paramètres objectifs. Nous pouvons penser que la réalité perceptive d'une dysphonie de paralysie laryngée relève plus de ces paramètres que de la raucité et de l'instabilité. Ces dimensions sont moins corrélées avec le grade, le souffle, et la sensation de faiblesse. Les mesures du rapport harmoniques/bruit seraient susceptibles de donner des informations objectives sur la présence de cette composante de bruit soufflé. Eskenazy et al (1990)^[77] n'ont pas trouvé de relation satisfaisante entre le signal/bruit et le souffle. Par contre Martin et al (1995)^[189] ont observé qu'une combinaison de variables acoustiques tels que le jitter, le shimmer, le rapport signal/bruit était plus pertinente pour une représentation objective du souffle qu'un paramètre pris isolément.

L'accord intra et inter-juges est relativement bon tant dans l'aptitude à discriminer le type de voix (pathologiques/témoins) que dans l'évaluation des critères de l'échelle G.I.R.B.A.S.. Nous notons cependant que le paramètre de forçage vocal n'est pas concrètement représenté chez tous les auditeurs. De même le paramètre de sensation de faiblesse ne recouvre pas la même réalité pour tous les juges puisqu'ils l'évaluent de manière très variable. Nous trouvons intéressant de recourir à une échelle d'évaluation à référents externes et de définir de manière plus précise les différents paramètres perceptifs pour augmenter la fiabilité des auditeurs. Les paramètres objectifs associés à l'apériodicité du signal vocal (Fo coef, Int coef, Jitt) semblent bien être liés aux paramètres subjectifs de l'échelle perceptuelle. Les paramètres de grade, souffle, et sensation de faiblesse sont les mieux corrélés. La combinaison des mesures objectives et l'introduction de la mesure du rapport signal/bruit pourraient apporter à ce sujet des renseignements supplémentaires.

5.6. Étude 7 : Application de la station E.V.A.2 à une population d'adultes âgés de 20 à 50 ans pour l'obtention de valeurs de référence. Comparaison avec un groupe de sujets féminins dysphoniques.

5.6.1. Introduction.

Notre étude s'inscrit dans le cadre de recherches visant à améliorer la pertinence et la fiabilité de l'évaluation vocale objective. Pour analyser nos échantillons nous avons choisi le logiciel E.V.A. 2 (évaluation vocale Assistée). Notre objectif est de fournir des valeurs de référence pour les paramètres acoustiques et aérodynamiques de cette station. Afin d'éprouver notre standard, nous avons comparé les résultats des sujets sains avec ceux des sujets féminins dysphoniques.

5.6.2. Matériel et méthode.

Notre échantillon est constitué de 88 sujets âgés de 20 à 50 ans, répartis en deux groupes, 48 femmes et 40 hommes. Ces sujets sont d'origines sociales et géographiques différentes. Les critères de sélection sont l'absence de pathologies laryngées et de pratique vocale. Nous avons demandé aux sujets trois émissions en vue de 3 analyses différentes. Un [a] simple est émis pour les mesures de hauteur tonale, d'intensité, de débit d'air buccal et pour les mesures de variabilité de ces mêmes paramètres. Un [a] est émis le plus longtemps possible pour les mesures du temps maximum phonatoire. Un [pa] est émis pour celles de pression sous-glottique. Certains paramètres tels que le fondamental, l'intensité sont repris pour les trois tâches, ce qui nous permettra de mesurer son effet sur ces paramètres.

Nous avons également comparé nos valeurs de référence à 15 sujets souffrant de pathologies telles que nodule, polype, sulcus, œdème de Reinke.

5.6.3. Résultats.

A. Statistiques descriptives.

Valeurs de référence pour les paramètres acoustiques et aérodynamiques de la station E.V.A.2 sur une population de sujets adultes âgés de 20 à 50 ans (n=88 ; F=48/ H=40).

Tableau 76 A : Statistique descriptive des mesures acoustiques et aérodynamiques pour la tâche : [a] simple.

[a] simple	Moyenne		DS*		Minimum		Maximum	
Sexes	F**	H***	F**	H***	F**	H***	F**	H***
Fondamental moy	216.57	115.86	24.18	16.76	180.06	85.27	295.10	172.51
Fo DS*	3.14	1.53	1.39	0.82	1.26	0.72	7.96	4.99
Fo C.V.	1.46	1.31	0.67	0.63	0.57	0.70	4.13	3.90
Jitter	2.68	0.61	2.08	0.23	0.99	0.23	13.17	1.28
RAP (x100)	8.08	2.96	6.97	1.17	0.002	0.001	0.05	0.007
Jitter %	1.26	0.52	1.05	0.18	0.40	0.22	6.83	1.09
Jitter ratio	12.56	5.22	10.48	1.76	3.96	2.20	68.38	10.90
Intensité moy	70.24	71.80	4.75	4.84	60.22	64.16	77.85	84.39
Int DS*	1.36	1.38	0.69	0.57	0.39	0.51	2.97	3.19
Int C.V.	1.95	1.94	1.01	0.86	0.54	0.65	4.66	4.55
Shimmer	0.33	0.32	0.14	0.20	0.11	0.09	0.64	0.88
APQ (x100)	6.35	0.10	2.53	4.41	0.02	0.04	0.15	0.20
Shimmer %	0.44	0.42	0.20	0.26	0.14	0.11	0.86	1.23

Légende : Moy : moyenne ; DS* : déviation standard ; F** : femmes ; H*** : Hommes.

[a] simple	Moyenne		DS		Minimum		Maximum	
	F	H	F	H	F	H	F	H
Fo	216.65	118.86	23.98	24.98	180.80	84.70	294.50	231.20
Sr	57.78	58.89	19.39	14.60	21.80	27.10	94.60	81.40
NNE	-4.44	-4.13	2.88	1.55	-12.71	-7.31	-1.07	-1.37
Rapport H/N	2.71	1.75	3.94	0.98	0.28	0.37	17.65	4.38
Sr HF	33.90	25.55	20.38	13.37	7.40	7.10	85.30	65.60
NNE Hf	-2.09	-1.36	1.79	0.87	-8.32	-4.63	-0.34	-0.32
H/N Hf	0.80	0.40	1.08	0.33	0.08	0.08	5.78	1.90
Fuite glo	1.84	2.32	0.96	0.99	0.19	0.33	4.19	4.12
F.G moy	0.13	0.17	6.92*	7.29*	0.01	0.02	0.30	0.31
FG sdt	1.02*	1.63*	8.30**	1.15*	0.001	0.004	0.05	0,06
FG C.V.	9.20	10.94	6.59	7.84	1.89	3.36	30.52	40.66

Tableau 76 B : Statistique descriptive du temps maximum phonatoire (TMP) pour la tâche [a] tenu.

[a] TMP	Moyenne		DS		Minimum		Maximum	
Sexes	F	H	F	H	F	H	F	H
TMP	15.16	20.34	4.88	6.60	6.61	10.89	27.52	39.57
Fo moyen	221.28	118.83	26.45	18.60	180.40	90.90	287.20	173.70
Intensité moy.	68.77	68.48	5.12	5.23	58.60	57.80	78.00	79.70
Volume exp.	1.76	2.76	1.00	1.66	0.01	0.09	4.79	6.68

Tableau 76 C : Statistique descriptive de la pression sous-glottique estimée pour la tâche [pa].

[pa] PSGE	Moyenne		DS		Minimum		Maximum	
Sexes	F*	H**	F	H	F	H	F	H
Pression	7.70	7.46	2.03	2.00	3.80	4.00	12.40	13.10
Intensité moy	70.29	74.40	7.73	4.06	36.10	65.30	83.10	82.40
Flux aérien moy	0.15	0.22	7.27	0.14	0.20	0.05	0.31	0.77

Légende : F : femmes ; H : hommes.

B. Effets de la répétition, du sexe, de la tâche en fonction des paramètres

Tableau 77 : Anova pour la fréquence fondamentale (Fo), l'intensité, et le flux en fonction du sexe, de la tâche et de la répétition.

E.V.A. 2		Femmes		Hommes	
Paramètres	Tâches	Test	Re-test	Test	Re-test
Fo	[a] simple	216.57	216.56	115.86	115.86
	[a] TMP	216.65	216.63	118.86	115.96
	[pa]	221.28	221.28	118.83	118.84
Intensité	[a] simple	70.23	70.24	71.8	71.81
	[a] TMP	68.77	68.77	68.46	68.47
	[pa]	70.28	70.29	74.4	74.4
Flux	[a] TMP	0.12	0.12	0.14	0.14
	[pa]	0.14	0.15	0.22	0.22

Nous avons effectué une ANOVA afin d'étudier la fréquence fondamentale en fonction du sexe, de la tâche et de la répétition (test/re-test).

L'effet principal est celui du sexe ($p < 0.0001$) : la fréquence fondamentale (Fo) est plus élevée chez la femme que chez l'homme. L'analyse ne montre pas d'effet principal de la répétition au test et au re-test ($p = 0.274$ N.S.). Enfin, nous notons un effet de la tâche ([a] simple, [a] tenu, [pa]) [Tableau 77].

Nous observons une interaction significative entre la tâche du [a] simple et du [pa] au niveau de la pression sous-glottique estimée ($p=0.002$). Le F_0 mesuré lors de la tâche du [pa] est supérieur à celui obtenu pour le [a] simple. Nous notons aussi une interaction significative entre la tâche [pa] et le [a] tenu ($p=0.013$) : la mesure du F_0 à la tâche [pa] est supérieure à celle que fournit le [a] tenu.

L'analyse de variance qui suit porte sur l'intensité moyenne (Int). Nous notons un effet du sexe ($p=0.0472$) et de la tâche ($p<0.0001$). Il n'y a pas d'effet de répétition. L'analyse montre d'ailleurs une interaction significative entre le sexe et la tâche ($p=0.004$). Nous observons une interaction entre le sexe et la tâche [pa] ($p=0.0006$). L'Int mesurée lors de la tâche du [pa] est plus élevée chez l'homme que chez la femme. Il y a aussi une interaction entre la tâche et le sexe masculin ($p<0.0001$). Chez l'homme uniquement, l'Int subit des modifications significatives lors de chacune des tâches. En effet, l'Int mesurée au [a] simple est supérieure à celle mesurée lors du [a] tenu ($p=0.0008$) et inférieure à celle mesurée lors du [pa] ($p=0.005$). De même, l'Int mesurée au [a] tenu est inférieure à celle obtenue au [pa] ($p<0.0001$).

Enfin, la troisième analyse concerne l'étude du flux. Ici encore, il y a un effet à la fois du sexe ($p=0.005$) et de la tâche ($p<0.0001$), et pas d'effet de répétition ($p=0.531$ N.S.).

Nous observons une interaction significative entre le sexe et la tâche ($p=0.022$), et une interaction significative du sexe et du flux mesuré à la tâche du [pa] ($p=0.0003$). Le flux mesuré lors du [pa] est plus élevé chez l'homme que chez la femme. En outre, il y a une interaction plus forte entre le sexe masculin et la tâche ($p<0.0001$) ; chez l'homme, la mesure du flux à la tâche du [pa] est plus élevée qu'à celle obtenue à la tâche du [a] tenu.

Nous avons ensuite effectué une ANOVA afin de connaître l'effet du sexe et de la répétition sur chacun de nos paramètres.

Nous observons un effet de répétition pour le F_0 SD ($p=0.008$), le F_0 C ($p=0.015$), l'Int Coef ($p=0.035$) et la FG Coef ($p=0.012$). Les valeurs obtenues au re-test sont supérieures à celles du test.

Nous notons plusieurs effets relatifs au sexe. Le F_0 SD est plus élevé chez les femmes ($p<0.001$). De même, les valeurs obtenues au Jitt ($p<0.001$), au RAP ($p<0.001$), au Jitt% ($p<0.001$) et au Jitt ratio ($p<0.001$) sont plus grandes chez elles.

Pour le paramètre APQ, les hommes obtiennent des mesures significativement inférieures à celles des femmes ($p < 0.001$).

Les paramètres de la fuite glottique (FG) ont des valeurs moins élevées chez les femmes que chez les hommes : pour FG ($p = 0.024$) ; pour FG Moy. ($p = 0.017$) ; pour FG SD ($p = 0.010$). En outre, les valeurs obtenues au Temps Maximum Phonatoire (TMP) et au Volume Expiratoire (VE) évoluent de la même manière (respectivement $p < 0.001$ et $p < 0.001$).

Les paramètres de bruit présentent des valeurs plus élevées pour les femmes que pour les hommes : pour SRHF, $p = 0.033$; pour NNEHF, $p = 0.024$; pour HNFHF, $p = 0.029$.

Enfin, nous avons évoqué la part de variabilité due au sujet par rapport à la variabilité globale. Nous notons, pour quatre paramètres, une erreur "relative" significative (Erreur Standard Moyenne (ESM) du Fo SD (1.47%), ESM du Fo Coef (1.52%), ESM de l'Int Coef (0.55%) et ESM du FG Coef (1.02%). Ces différences sont toutefois très faibles et la variabilité au test/re-test quasi inexistante. En effet, la corrélation intra-classe pour ces quatre paramètres est supérieure à 0.99.

- **Résultats concernant la population de sujets dysphoniques.**

Rappelons que notre échantillon de sujets dysphoniques est inséré à titre d'exemple. Il se compose de 12 femmes et 3 hommes. L'étiologie de la dysphonie consiste dans 8 cas en un nodule, dans 2 cas un sulcus, dans 4 cas un œdème de Reinke, dans un cas un polype. Cet effectif est restreint et ne peut être considéré comme représentatif de sa population. Néanmoins, il nous paraissait intéressant d'illustrer les valeurs des paramètres acoustiques et aérodynamiques dans une population de sujets dysphoniques.

C. Statistique descriptive pour toutes les pathologies confondues.

Nous avons effectué une statistique descriptive (moyenne, écart-type, minimum, maximum) sur la population de 15 sujets dysphoniques âgés de 20 à 50 ans, pour chacune des trois tâches. Cette statistique descriptive est réalisée pour toutes les pathologies confondues.

Tableau 78 A : Statistique descriptive des mesures acoustiques et aérodynamiques, pour la tâche : [a] simple toutes pathologies confondues.

[a] simple	Moyenne		DS		Minimum		Maximum	
Sexes	F*	H**	F	H	F	H	F	H
F0. moy	191.12	133.76	35.38	7.53	140.65	125.2	265.02	139.4
F0 déviation sdt	2.5	2.43	1.18	2.17	1.01	1.13	4.65	4.94
F0 C.V.	1.39	1.8	0.8	1.57	0.48	0.89	2.98	3.62
Jitter	1.5	0.67	0.6	0.34	0.8	0.4	2.9	1.05
RAP	0.005	0.003	0.002	0.001	0.002	0.002	0.01	0.004
Jitter %	0.81	0.5	0.4	0.24	0.31	0.32	1.53	0.77
Jitter ratio	8.12	4.8	3.9	2.16	3.15	3.16	15.27	7.25
Intensité moy	71.55	77.27	7.37	1.75	59.45	75.95	85.05	79.25
Int déviation sdt	1.77	1.63	0.96	0.34	0.35	1.26	3.14	1.94
Int C.V.	2.53	2.11	1.38	0.48	0.44	1.59	4.4	2.54
Shimmer	0.46	0.52	0.24	0.39	0.09	0.23	0.941	0.954
APQ	0.08	0.11	0.04	0.05	0.02	0.06	0.16	0.16
Shimmer %	0.61	0.63	0.36	0.5	0.1	0.27	1.42	1.2

Légende : Moy : moyenne ; F* : femmes ; H** : hommes.

[a] simple	Moyenne		DS		Minimum		Maximum	
Sexes	F	H	F	H	F	H	F	H
F0	191.23	133.23	35.28	7.2	141.1	125.1	264.7	138.8
Sr	56.69	63.17	14.31	6.2	34.1	56	76.1	66.9
NNE	-3.85	-4.38	1.44	0.7	-6.21	-4.8	-1.81	-3.57
Rapport H/N	1.55	1.76	0.84	0.43	0.52	1.27	3.18	2.02
Sr hf	23.68	24.63	16.82	17.56	4.4	4.4	63.6	35.9
NNE hf	-1.3	-1.3	1.17	0.96	-4.39	-1.93	-0.2	-0.2
H/N hf	0.4	0.37	0.47	0.28	0.05	0.05	1.75	0.56
Fuite glo	3.28	3.45	1.9	1.95	0.34	1.48	6.26	5.37
Fuite glo moy	0.24	0.27	0.13	0.15	0.023	0.114	0.39	0.408
Fuite glo sdt	0.014	0.017	0.009	0.014	0.005	0.008	0.037	0.033
Fuite glo C.V.	12.28	6.2	20.06	2.35	3.1	3.54	73.92	7.99

Légende : Moy : moyenne ; F* : femmes ; H** : hommes.

Tableau 78 B : Statistique descriptive du temps maximum phonatoire (TMP) à la tâche : [a] tenu toutes pathologies confondues.

[a] TMP	Moyenne		DS		Minimum		Maximum	
Sexes	F*	**H	F	H	F	H	F	H
TMP	9.84	8.16	2.44	4.42	6.46	4.2	14.65	13
F0 moy.	184.4	139.33	36.14	7.63	126.7	130.8	239.7	145.5
Intensité moy.	68.04	77.13	5.39	4.93	57.5	73.3	76.7	82.7
Volume exp.	2.14	2.12	0.64	1.2	0.56	1.36	2.9	3.51
Flux aérien moy	0.22	0.27	0.069	0.06	0.067	0.2	0.36	0.32

Légende : Moy : moyenne ; F* : femmes, H** : hommes.

Tableau 78 C : Statistique descriptive de la pression sous-glottique estimée pour la tâche: [pa] toutes pathologies confondues.

[pa] PSGE	Moyenne		DS		Minimum		Maximum	
Sexes	F*	H**	F	H	F	H	F	H
Pression	10.37	15.4	2.41	4.08	7.2	12.2	14.9	20
Intensité moy	73.04	76.73	7.56	7.29	58	70.6	83.5	84.8
Flux aérien moy	0.24	0.34	0.08	0.17	0.11	0.15	0.36	0.49

Légende : MOY : moyenne ; F* : femmes, H** : hommes.

Nous n'avons pas pu effectuer une analyse de variance (ANOVA) afin d'étudier l'effet du sexe sur les différents paramètres acoustiques et aérodynamiques, étant donné le nombre restreint de sujets masculins dysphoniques (n=3).

C 1 Statistique descriptive par pathologie.

Nous avons choisi d'illustrer la statistique descriptive (moyenne, écart-type, minimum, maximum) pour les sujets féminins présentant un nodule ou un œdème de Reinke. Le choix s'est porté sur ces pathologies, parce qu'elles étaient les seules à se présenter plusieurs fois dans notre effectif de sujets féminins (nodule n=8. F=7/H=1 ; œdème de Reinke n=4. F=4/H=0 ; Sulcus n=2. F=1/H=1 ; polype n=1. F=0/H=1).

Statistique descriptive pour les 7 sujets féminins présentant un NODULE

Tableau 79 A : Statistique descriptive des mesures acoustiques et aérodynamiques pour la tâche : [a] simple pour les sujets féminins porteurs de nodules.

[a] simple	Moyenne	DS	Minimum	Maximum
Fondamental moy	203.07	34.89	152.31	265.02
F0 déviation std	2.04	0.74	1.01	3.1
F0 C.V.	1.05	0.47	0.48	1.67
Jitter	1.54	0.74	0.84	2.87
RAP	0.005	0.003	0.002	0.01
Jitter %	0.81	0.46	0.31	1.53
Jitter ratio	8.047	4.63	3.15	15.27
Intensité moy	71.49	9.29	59.45	85.05
Int déviation std	1.6	0.97	0.35	3.11
Int C.V.	2.33	1.42	0.44	4.11
Shimmer	0.47	0.28	0.091	0.94
APQ	0.075	0.045	0.02	0.16
Shimmer %	0.64	0.44	0.1	1.42

[a] simple	Moyenne	DS	Minimum	Maximum
F0	202.93	34.86	153.3	264.7
Sr	65.73	7.06	56.5	76.1
NNE	-4.73	0.92	-6.21	-3.61
Rapport H/N	2.03	0.66	1.3	3.18
Sr hf	26.84	21.4	4.4	63.6
NNE hf	-1.55	1.49	-4.39	-0.2
H/N hf	0.51	0.6	0.05	1.75
Fuite glo	3.47	1.98	0.38	6.26
Fuite glo moy	0.25	0.13	0.02	0.4
Fuite glo sdt	0.02	0.01	0.008	0.04
Fuite glo C.V.	15.24	25.98	3.1	73.92

Tableau 79 B : Statistique descriptive du temps maximum phonatoire pour la tâche : [a] tenu pour les sujets féminins porteurs de nodules.

[a] TMP	Moyenne	DS	Minimum	Maximum
TMP	10.56	2.77	6.46	14.65
F0 moyen	197.93	24.4	169.5	239.7
Intensité moyen	67.43	6.46	57.5	76.7
Volume exp.	2.35	0.52	1.32	2.9
Flux aérien moyen	0.23	0.056	0.155	0.34

Tableau 79 C : Statistique descriptive de la pression sous-glottique estimée pour la tâche : [pa] pour les sujets féminins porteurs de nodules.

[pa] PSGE	Moyenne	DS	Minimum	Maximum
Pression	10.71	3.03	7.2	14.9
Intensité moy	73.94	8.84	58	83.5
Flux aérien moy	0.21	0.11	0.021	0.35

Statistique descriptive pour les 4 sujets féminins présentant un OEDEME DE REINKE

Tableau 80 A : Statistique descriptive pour les mesures acoustiques et aérodynamiques pour la tâche : [a] simple pour les sujets féminins porteurs d'œdème de Reinke.

[a] simple	Moyenne	DS	Minimum	Maximum
Fondamental moy	167.69	32.48	140.65	214.87
F0 déviation sdt	3.56	1.29	1.83	4.65
F0 C.V.	2.15	0.85	1.3	2.98
Jitter	1.32	0.37	0.83	1.71
RAP	0.005	0.002	0.002	0.007
Jitter %	0.84	0.34	0.39	1.22
Jitter ratio	8.48	3.47	3.86	12.17
Intensité moy	72.73	4.28	68.29	78.52
Int déviation sdt	2.25	0.97	0.88	3.14
Int C.V.	3.14	1.4	1.12	4.4
Shimmer	0.45	0.25	0.3	0.82
APQ	0.11	0.043	0.05	0.14
Shimmer %	0.58	0.31	0.37	1.03
[a] simple	Moyenne	DS	Minimum	Maximum
F0	168.23	32.77	141.1	215.9
Sr	38.88	3.95	34.1	43.5
NNE	-2.145	0.28	-2.48	-1.81
Rapport H/N	0.64	0.11	0.52	0.77
Sr hf	17.53	6.68	8.7	24.7
NNE hf	-0.85	0.35	-1.23	-0.4
H/N hf	0.22	0.095	0.1	0.33
Fuite glo	3.4	2.08	0.34	4.95
Fuite glo moy	0.25	0.16	0.023	0.39
Fuite glo sdt	0.012	0.006	0.005	0.018
Fuite glo C.V.	8.64	8.63	3.61	21.53

Tableau 80 B : Statistique descriptive du temps maximum phonatoire pour la tâche : [a] tenu pour les sujets féminins porteurs d'œdème de Reinke.

[a] TMP	Moyenne	DS	Minimum	Maximum
TMP	8.95	1.89	6.89	11.27
F0 moyen	152.05	37.06	126.7	206.6
Intensité moyen	69.9	3.77	66.8	75.3
Volume exp.	2.15	0.072	2.072	2.23
Flux aérien moy	0.25	0.05	0.2	0.3

Tableau 80 C : Statistique descriptive de la pression sous-glottique estimée pour la tâche : [pa] pour les sujets féminins porteurs d'œdème de Reinke.

[pa] PSGE	Moyenne	DS	Minimum	Maximum
Pression	9.7	1.45	8.4	11.4
Intensité moy	71.38	6.92	65.3	81.2
Flux aérien moy	0.26	0.1	0.11	0.36

- **Comparaison des sujets sains et des sujets dysphoniques féminins**

Comme nous l'avons déjà mentionné, une comparaison entre les sujets sains et dysphoniques est difficile étant donné l'effectif différent des 2 populations (88 sujets sains et 15 dysphoniques). Cependant, le nombre de sujets féminins dysphoniques est plus élevé que celui de sujets masculins dysphoniques . Aussi, avons-nous sélectionné les sujets féminins dysphoniques présentant un nodule ou un œdème de Reinke. Dans notre effectif de 48 sujets féminins sains, nous avons sélectionné 10 sujets en utilisant une table de nombres aléatoires, qui nous permettait de choisir un sous-échantillon de sujets féminins sains représentatif de l'échantillon de départ. Ce nouvel échantillon de sujets choisis au hasard nous permettait de comparer les deux populations. Précisons cependant que cette statistique comparative n'a d'autre prétention que d'être illustrative et qu'il convient de rester prudent par rapport à toute conclusion.

C.2. Statistique descriptive des sujets féminins sains sélectionnés.

Nous avons effectué une statistique descriptive (moyenne, écart-type, minimum, maximum) pour les 10 sujets féminins sains sélectionnés.

Tableau 81 A : Statistique descriptive pour les mesures acoustiques et aérodynamiques pour la tâche : [a] simple pour les sujets féminins sains.

[a] simple	Moyenne	DS	Minimum	Maximum
Fondamental moy	213.14	20.05	184.68	250.68
F0 déviation sdt	2.98	0.9	1.89	4.24
F0 C.V.	1.4	0.42	0.97	2.15
Jitter	2.31	1.47	1.17	6.31
RAP	0.007	0.005	0.004	0.02
Jitter %	1.01	0.73	0.59	3.07
Jitter ratio	11	7.23	5.93	30.66
Intensité moy	70.72	4.27	61.43	74.88
Int déviation sdt	1.21	0.51	0.56	2.38
Int C.V.	1.73	0.77	0.75	3.4
Shimmer	0.31	0.11	0.2	0.58
APQ	0.06	0.02	0.04	0.08
Shimmer %	0.4	0.15	0.27	0.78

[a] simple	Moyenne	DS	Minimum	Maximum
F0	213.3	20.42	184.6	252.6
Sr	62.76	20.07	21.8	92.1
NNE	-5.04	2.98	-11.04	-1.07
Rapport H/N	3.08	3.57	0.28	11.7
Sr hf	37.74	23.51	10.3	78
NNE hf	-2.41	1.99	-6.58	-0.47
H/N hf	0.94	1.07	0.11	3.55
Fuite glo	1.56	0.9	0.31	2.62
Fuite glo moy	0.11	0.067	0.02	0.2
Fuite glo sdt	0.01	0.01	0.001	0.05
Fuite glo C.V.	10.86	7.11	3.32	25.16

Tableau 81 B : Statistique descriptive du temps maximum phonatoire pour la tâche : [a] tenu pour les sujets féminins sains.

[a] TMP	Moyenne	DS	Minimum	Maximum
TMP	15.23	4.82	7.08	23.76
F0 moyen	219.4	25.24	193.1	273.6
Intensité moyen	68.63	5.11	59.9	74.8
Volume exp.	1.68	0.96	0.03	2.77
Flux aérien moy	0.12	0.07	0.002	0.22

Tableau 81 C : Statistique descriptive de la pression sous-glottique estimée pour la tâche : [pa] pour les sujets féminins sains.

[pa] PSGE	Moyenne	DS	Minimum	Maximum
Pression	7.8	2.39	5.1	12.4
Intensité moy	71.98	6.55	62.7	83.1
Flux aérien moy	0.14	0.09	0.04	0.31

B. Effet de la tâche et de la présence ou non d'une dysphonie en fonction des paramètres.

Tableau 82 : Comparaison des mesures de Fo, intensité et flux entre les 3 groupes de sujets féminins.

E.V.A. 2	Femmes			
Paramètres	Tâches	Nodules	Œdèmes de R.	Sains
F0	[a] simple	203.07	167.69	213.14
	[a] TMP	197.93	152.05	219.4
Intensité	[a] simple	71.49	72.73	70.72
	[a] TMP	67.43	69.9	68.63
	[pa] PSGE	73.94	71.38	71.98
Flux	[a] TMP	0.23	0.25	0.12
	[pa] PSGE	0.21	0.26	0.14

Afin d'étudier la fréquence fondamentale (F0) en fonction de la tâche et de la présence ou non d'une dysphonie nous avons effectué une ANOVA.

L'analyse montre un effet de la tâche. Le Fo du [a] simple apparaît plus élevé que le Fo du [a] tenu chez les sujets présentant un nodule ($p=0.036$) et ceux

présentant un œdème de Reinke ($p=0.001$). Par contre, le F_0 du [a] tenu apparaît plus élevé que le F_0 du [a] simple chez les sujets sains ($p=0.042$).

Il y a aussi un effet de la dysphonie ($p=0.01$). Le F_0 du [a] simple ($p=0.04$) et le F_0 du [a] tenu ($p=0.002$) apparaissent tous deux significativement plus élevés chez les sujets sains.

La comparaison des moyennes à l'aide du Test de Différence Significative Honnête de Tukey montre que les sujets présentant un œdème de Reinke se différencient significativement des sujets sains ($p=0.007$) et non significativement des sujets porteurs de nodule(s) ($p=0.07$ N.S.).

Afin d'étudier l'intensité en fonction de la tâche et de la présence ou non d'une dysphonie nous avons effectué une ANOVA.

L'analyse montre un effet significatif de la tâche uniquement chez les sujets présentant un nodule ($p=0.02$). L'intensité du [a] tenu y apparaît significativement moins élevée ($p=0.007$).

Il n'y a pas d'effet de la dysphonie (pour la tâche du [a] simple ($p=0.86$ N.S.) ; [a] tenu ($p=0.76$ N.S.) ; [pa] TMP ($p=0.87$ N.S.)).

Nous avons aussi effectué une ANOVA afin d'étudier le flux moyen en fonction de la tâche et de la présence ou non d'une dysphonie.

Nous notons un effet principal de la dysphonie ($p=0.008$) et aucun effet de la tâche ($p=0.35$ N.S.).

Le flux moyen apparaît plus élevé chez les sujets dysphoniques pour la tâche du [a] simple ($p=0.001$) et pour la tâche du [a] tenu ($p=0.04$).

La comparaison des moyennes à l'aide du Test de Différence Significative Honnête de Tukey montre que les sujets porteurs de nodules se différencient significativement des sujets sains ($p=0.007$) et non significativement des sujets présentant un œdème de Reinke ($p=0.9$ N.S.). De même, les sujets présentant un œdème de Reinke se différencient significativement des sujets sains ($p=0.01$).

Nous avons effectué une ANOVA afin de connaître l'effet de la dysphonie sur chacun des autres paramètres acoustiques et aérodynamiques.

Nous ne mentionnerons que les valeurs significatives.

Tableau 83 : Comparaison des mesures de fuite glottique, de Sr, de temps maximum phonatoire, d' APQ et de pression sous-glottique estimée entre les 3 groupes de sujets féminins.

E.V.A. 2	Femmes			
Paramètres	Tâches	Nodules	Œdèmes de R.	Sains
Fuite glottique	[a] simple	3.47	3.4	1.56
Fuite glot. Moy	[a] simple	0.25	0.25	0.11
Sr	[a] simple	65.73	38.88	62.76
TMP	[a] tenu	10.56	8.95	15.22
APQ	[a] simple	0.075	0.11	0.06
Pression	[pa] PSGE	10.71	9.7	7.8

Nous observons un effet de la dysphonie pour la fuite glottique ($p=0.04$), la fuite glottique moyenne ($p=0.03$), et la pression ($p=0.002$). Toutes ces valeurs apparaissent significativement plus élevées chez les sujets dysphoniques.

Il y a un effet sur le TMP ($p=0.01$), qui apparaît moins élevé chez les sujets dysphoniques.

Les sujets présentant un œdème de Reinke ont des valeurs pour Sr significativement plus basses ($p=0.02$) que les sujets sains ($p=0.049$) et les sujets présentant un nodule ($p=0.04$).

L'effet de la dysphonie sur l'APQ est non significatif ($p=0.056$).

Comparaison TMP sujets féminins dysphoniques et sains

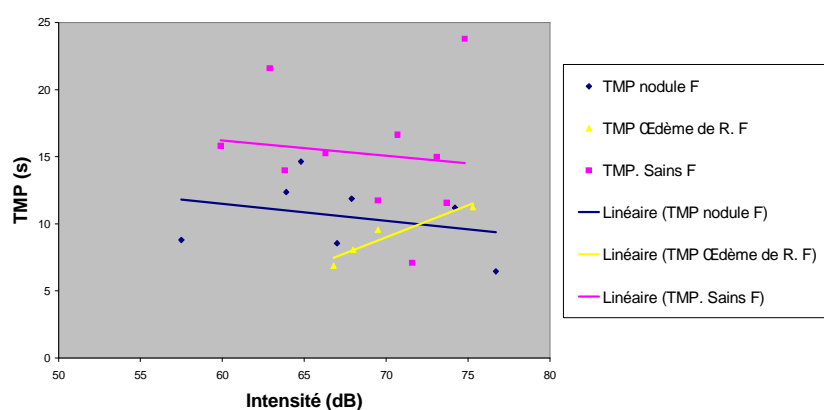


Figure 32 : Comparaison entre les sujets dysphoniques et les sujets sains du temps maximum phonatoire.

Comparaison fuite glo. sujets féminins dysphoniques et sains

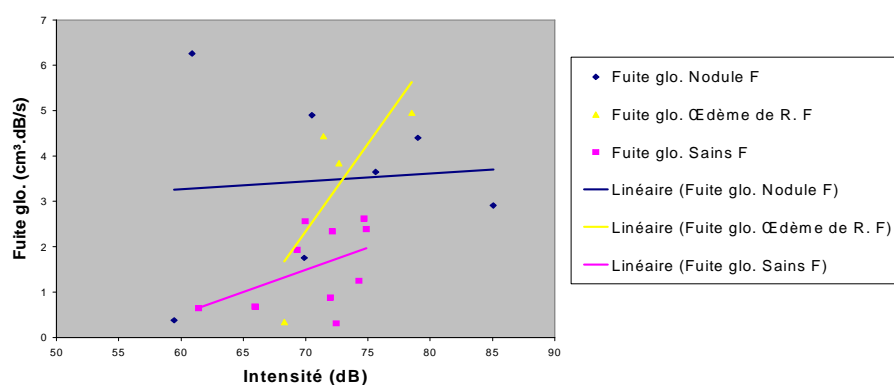


Figure 33 : Comparaison entre sujets dysphoniques et sujets sains de la fuite

Comparaison PSGE sujets féminins dysphoniques et sains

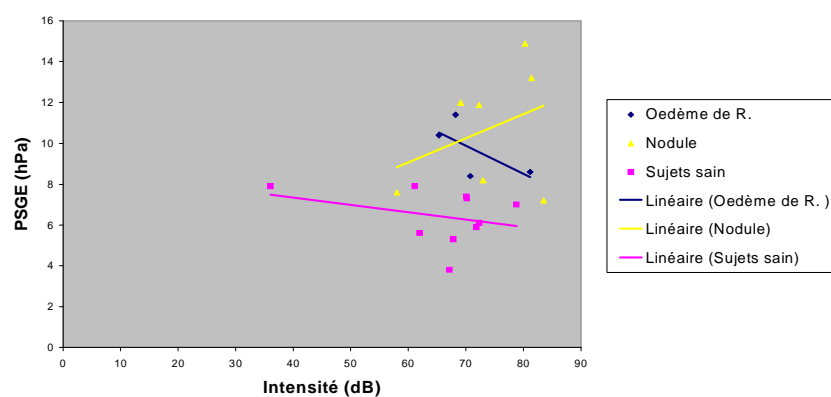


Figure 34 : Comparaison entre les sujets dysphoniques et les sujets sains de la PSGE.

- Il apparaît que la mesure du **fondamental** est plus élevée à la tâche du [a] simple qu'à la tâche du [a] tenu chez les sujets dysphoniques, la situation inverse est observée chez les sujets sains.
- Le **fondamental** est plus élevé chez les sujets sains.
- L'**intensité** mesurée lors du [a] tenu apparaît significativement moins élevée chez les sujets présentant un nodule.
- Les mesures du **flux**, de la **fuite glottique**, de la **fuite glottique moyenne** et de la **pression sous-glottique** estimée apparaissent plus élevées chez les sujets dysphoniques.
- Le **temps maximum phonatoire** est plus élevé chez les sujets sains.
- Les valeurs de **Sr** (énergie en pour-cent du fondamental et des harmoniques par rapport à l'énergie totale) sont significativement moins élevées chez les sujets présentant un œdème de Reinke.

5.6.4. Discussion et conclusion.

Nos résultats confirment la fiabilité du logiciel E.V.A. 2^[296]. Notre choix s'est porté sur une fenêtre d'analyse d'une seconde, 100 millisecondes après le début du voisement. De prime abord aucune différence n'est mise en évidence au test re- test. Cependant, une analyse statistique a révélé d'infimes différences entre les deux analyses sur 4 paramètres : la déviation standard de la fréquence fondamentale, le coefficient de variation de la fréquence fondamentale et de l'intensité et la fuite glottique. Ce léger manque de reproductibilité mérite d'être exploré sur une population de sujets plus étendue. L'expérimentateur ne peut être mis en cause puisqu'il positionne la fenêtre d'analyse en se basant sur la courbe du fondamental et que simultanément, l'analyse s'effectue sur toutes les courbes. Une erreur de l'expérimentateur aurait fait varier les valeurs de tous les paramètres. Le calibrage des instruments de mesure se fait automatiquement. Il est possible que d'un jour à l'autre, la variation de température ambiante induise de minimes variations se répercutant sur l'analyse. Le re-test s'est effectué un mois plus tard.

Le choix d'analyser la partie stable a déjà été discutée précédemment.

Nos résultats statistiques concernant la fréquence fondamentale montrent la différence attendue entre les deux sexes. Dans la tâche du [a] simple, le Fo moyen est de 216.57 Hz pour les femmes et 115.86 Hz pour les hommes. Dans la tâche du [a] tenu, le Fo moyen est de 221.28 Hz pour les femmes et de 118.83 Hz pour les hommes.

Nos résultats montrent un effet significatif de la tâche. En effet, le fondamental du [a] tenu est plus élevé que celui du [a] simple. Ces résultats peuvent s'expliquer par un comportement d'adaptation du sujet par rapport à la tâche. Il lui était demandé de produire un [a] tenu le plus longtemps possible. Face à cette consigne le sujet peut chercher une stratégie lui permettant d'allonger la durée de son émission vocale. Pour ce faire, il augmente sa fréquence fondamentale et régule son intensité afin d'économiser son souffle. Ces résultats corroborent ceux de Raes (1996)^[242] et Heylen (1996)^[118].

L'intensité n'est pas significativement différente entre les hommes et les femmes (70.24 dB et 71.8 dB respectivement). Nos résultats sont similaires à ceux de Gelfer et al (1997)^[89]. Les auteurs n'ont rapporté aucune différence. Ses résultats en conversation spontanée étaient de 70.42 dB pour les hommes et 68.15 dB pour les femmes. Cependant, il apparaît que la mesure est significativement plus élevée chez les hommes pour la tâche du [pa] (74.4 dB). Le fait que l'intensité ne subit des modifications que chez les hommes peut être également expliqué par un comportement d'ajustement à la tâche. Selon Boone et al (1994)^[26], l'intensité peut varier en fonction de la tâche demandée au sujet. Il en va de même pour le discours, l'orateur ajuste son niveau d'intensité non seulement en fonction de la distance par rapport à son interlocuteur mais aussi par rapport au nombre d'interlocuteurs. Nos résultats ne corroborent pas l'hypothèse de Baken et al (1987)^[14] qui explique que la difficulté majeure dans la détermination de l'intensité est d'éliminer l'influence des moments de pause et d'intervalle que peut présenter un discours tel que [pa], [pa], [pa]. Ces moments de pause, intégrés dans la mesure du signal peuvent réduire la mesure d'intensité moyenne.

Les valeurs du flux moyen mesuré lors de la tâche du [pa] sont plus élevées chez l'homme que chez la femme. Ces résultats similaires à ceux obtenus pour l'intensité sont liés.

Les mesures du temps maximum phonatoire, du volume expiré et de la fuite glottique sont également plus élevées chez l'homme, quelle que soit la tâche. Cela s'explique simplement par une capacité vitale supérieure chez l'homme. Une étude de Zajac et al (1996)^[359] a comparé 42 sujets (21 Hommes, 21 femmes). Les résultats montrent que les hommes ont un flux d'air plus élevé que les femmes (149 ml/s chez les hommes versus 129 ml/s chez les femmes). Les auteurs concluent qu'une plus grande élasticité pulmonaire serait à l'origine de cette différence. Cependant, il faut accueillir cette opinion avec la plus grande prudence. Les valeurs de jitter apparaissent plus élevées chez la femme que chez l'homme. Selon Baken^[14] dans la littérature^[131. 138. 166. 211. 243] les valeurs concernant le jitter présentent une grande hétérogénéité. La valeur standard qu'il propose est de 0.37%. Nos résultats corroborent ceux de Giovanni et al (1996)^[95] en ce qui concerne la disparité hommes/ femmes. Par contre notre valeur moyenne est haute (1.26 pour les femmes et 0.52 pour les hommes). Néanmoins, nos résultats se maintiennent sous la valeur des 5% considérée comme fiable par Titze (1995)^[328]. Les conseils de Titze au sujet du jitter restent d'actualité. Ce paramètre doit être utilisé avec prudence en pathologie.

Les résultats concernant la comparaison avec les sujets dysphoniques féminins montrent que la fréquence fondamentale est plus élevée à la tâche du [a] simple qu'à celle du [a] tenu chez les sujets dysphoniques. Ceux-ci montrent un comportement inverse aux sujets sains pour la tâche du TMP. Giovanni 1998^[98] explique que les différences observées dans les forces exercées sur la glotte et les différences de fonctionnement sont dues aux diverses irrégularités de la géométrie des tissus et à leurs propriétés mécaniques. A fortiori, ces divergences s'accroissent dans le cadre de pathologies telles que les nodules, les kystes ou les polypes. Dès lors, les différences de contraintes biomécaniques sont plus importantes ainsi que les perturbations du mouvement mécanique qui en découlent. La pathologie accentue l'asymétrie des cordes vocales et contribue d'autant plus à la perturbation de l'oscillation. Les cordes sont désynchronisées. Les sujets tentent de compenser leur difficulté vocale et de rentabiliser au mieux la tâche du temps maximum phonatoire, mais en vain.

Le fait que la fréquence fondamentale apparaisse moins élevée chez les sujets dysphoniques montre la différence que faisait attendre la littérature entre les sujets sains et dysphoniques^[154].

Les mesures de flux, de la fuite glottique, et de la pression sont plus élevées chez les sujets dysphoniques. Ce phénomène refléterait un moyen compensatoire pour maintenir une pression sous-glottique constante en présence d'un flux d'air glottique plus élevé que la normale. En effet, le locuteur pour maintenir l'efficacité, la portée et l'intelligibilité de sa voix, peut augmenter son souffle expiratoire. Hélas, cette manœuvre ne pourra pas compenser le déficit.

Chez les sujets sains, comme chez les sujets dysphoniques, nos résultats concernant la PSGE montrent des valeurs corroborant celles de la littérature. Il est toutefois intéressant de mentionner qu'il est toujours difficile de comparer des résultats en valeur absolue en raison des nombreuses sources de variabilité des mesures effectuées. En particulier, notre méthode ne fait pas référence à la position exacte de la sonde dans la cavité buccale à la différence d'autres études qui en font un élément déterminant de la précision de la méthode. Baken^[14], par exemple, a développé un dispositif de maintien de la sonde introduite dans la cavité buccale, à l'aide d'un moulage des canines. Dans le cadre de nos expériences, un tel dispositif n'était pas nécessaire étant donné les caractéristiques de la sonde. Celle-ci était suffisamment rigide pour ne pas être écrasée entre les dents et elle présentait en son extrémité plusieurs orifices empêchant l'obstruction par la salive. Outre ces considérations d'ordre technique d'une part, la cohérence d'une PSGE anormalement augmentée chez les sujets dysphoniques d'autre part justifient suffisamment le procédé choisi. Les aspects biomécaniques des cordes vocales^[98], l'état psychologique du sujet dysphonique^[34], et les mouvements compensatoires inadéquats rendent la sonorisation difficile et ce malgré une augmentation des valeurs de ce paramètre.

Chapitre 6

Index de la qualité vocale.

Ce chapitre est consacré à l'index de la qualité vocale. Il s'agit d'un outil qui réduit les mesures acoustiques et aérodynamiques à une seule valeur. L'intérêt de ces index est son aspect pratique étayé sur des fondements fiables et valides.

6.1. L'index de Piccirillo.

Piccirillo et al (1998)^[230] ont tenté, dans un premier temps, d'identifier les paramètres vocaux importants, et dans un second temps^[232], de créer un indice vocal multidimensionnel combinant les paramètres pertinents pour évaluer la fonction vocale.

Trois indices sont créés sur base de l'évaluation de 132 sujets répartis en deux groupes (97 sujets dysphoniques, 34 sujets avec troubles bénins, 54 avec troubles malins, et 9 laryngectomisés : 35 autres sont volontaires sans plainte particulière). Le premier indice proposé est basé sur 14 paramètres vocaux incluant des mesures acoustiques (intensité vocale, fréquence, étendue fréquentielle), des mesures aérodynamiques (pression sous-glottique, flux d'air, temps maximum phonatoire) et une mesure électroglottographique (taux de changement des zones de contact). Le deuxième indice est basé sur le score du rapport de probabilité d'occurrence de chacun de ces paramètres. Le troisième est dérivé des résultats de l'analyse statistique. A l'analyse de régression, seuls 4 paramètres sur 14 semblent être suffisants. Ces 4 paramètres correspondent à la fréquence, au flux d'air aux lèvres, au temps maximum phonatoire, et à la pression sous-glottique. Les auteurs précisent que leur indice est plutôt un instrument de classification des dysphonies qu'un instrument de prédiction ou d'efficacité du traitement. Une seconde étude^[232] sur 33 patients et 12 sujets sains est réalisée. Les résultats montrent que plus le score à l'index de probabilité d'occurrence des paramètres vocaux est élevé, plus la probabilité que le patient soit dysphonique augmente. A l'index de régression, plus la valeur de l'index est bas, plus il est attendu que le patient soit dysphonique. La corrélation entre l'index du rapport de probabilité d'occurrence et l'échelle G.R.B.A.S. est statistiquement significative pour chaque paramètre. Lorsque l'index augmente, les scores de l'échelle augmentent aussi. La corrélation entre l'index de régression et les résultats de l'échelle perceptuelle montre une corrélation de 0.49 pour la raucité et 0.73 pour le souffle. L'auteur conclut que l'index du rapport de probabilité d'occurrence et l'index de régression sont satisfaisants. Ces deux indices peuvent servir de mesures objectives de la sévérité des troubles vocaux en pré- traitement. Cependant, Piccirillo et al

regrettent le nombre limité de sujets, le peu de diversité dans les troubles observés et l'absence d'évaluation post thérapeutique ce qui remet en question la sensibilité de l'index.

6.2. Le "Dysphonia Severity Index" de wuyts.

Le D.S.I. ou index de sévérité de la dysphonie est une mesure élaborée par Wuyts et al (1998)^[355]. Il s'agit d'une mesure objective de la qualité vocale basée sur une approche multiparamétrique. Il repose sur la combinaison pondérée de mesures vocales telles que la fréquence fondamentale la plus haute, l'intensité la plus basse, le temps maximum phonatoire, et le jitter en %. Il est construit sur l'équation suivante:

D.S.I.: $0.13 \times \text{TMP} + 0.0053 \times \text{Fo le plus haut} - 0.26 \times \text{Int la plus basse} - 1.18 \times \text{Jitter \%} + 12.4.$

Le choix des paramètres inclus dans le calcul du D.S.I. ont été déterminés par une procédure "stepwise logistic regression ". L'emploi de la fréquence fondamentale la plus élevée est justifié disent les auteurs par le fait que dans plus de 50% des cas de dysphonie, les cordes vocales ont un excès de masse. Cette masse présente sur la corde gêne la fréquence de vibration la plus haute. En conséquence, la fréquence fondamentale la plus élevée décroît. Les patients dysphoniques augmentent la résistance de la glotte, ils utilisent dès lors une plus grande pression pour la mise en vibration des cordes vocales et le maintien de cette vibration. Il s'ensuit une diminution de l'intensité la plus basse. Le jitter est une variable reflétant le degré d'irrégularité de la corde vocale dans une certaine limite. Le temps maximum phonatoire reflète l'efficacité de plusieurs mécanismes nécessaires à la production vocale tels que la pression sous-glottique, la résistance du flux d'air, la fermeture des cordes vocales. Ces quatre variables, fait remarquer l'auteur, sont toutes des indicateurs clairs de la dysphonie.

Les voix normales obtiennent un index de +5, les voix dysphoniques sévères, un index de -5. Plus le sujet est dysphonique, au plus son index est négatif.

Cet index est le reflet du caractère multidimensionnel de la voix. Il est reproductible et corrélé avec l'opinion du clinicien.

L'index a été éprouvé sur une population de 387 sujets de 18 à 80 ans provenant de six laboratoires (Base de données du B.S.G.V.D.). 53% sont des

femmes, 47 % des hommes ; 68 sujets constituent le groupe contrôle. Les résultats montrent une forte corrélation entre le D.S.I. et l'échelle perceptuelle G.R.B.A.S. (0.996), ainsi qu'avec le V.H.I. (Voice Handicap Index : [annexe 5]) (-0.79). Le D.S.I. traduit un bon accord entre la perception du clinicien (évaluation du G de l'échelle G.R.B.A.S.) et celle du patient lui-même (évaluation du V.H.I.). Ces corrélations significatives renforcent la validité de l'index.

La pratique clinique et le domaine de la recherche s'enrichissent d'un nouvel outil. La mesure des paramètres inclus dans le D.S.I. est accessible au laboratoire mais également au clinicien individuel. L'auteur souligne que la marge d'erreur est très faible, de l'ordre de 6%. La base de données élaborée par 6 laboratoires renforce la fiabilité de l'index. L'outil permet d'évaluer l'efficacité thérapeutique, de comparer diverses formes de thérapie, ainsi que les traitements chirurgicaux.

Les améliorations apportées au D.S.I. viseraient dans un premier temps les variables acoustiques qui seraient basées sur du langage spontané plutôt que sur la voyelle. Dans un second temps, les mesures de flux d'air et de pression sous-glottique seraient intégrées à l'index.

Le D.S.I. complète donc l'évaluation du sujet dysphonique. De Bodt (1997) ^[52] décrit dès lors l'évaluation sous forme de 4 composantes: l'inspection laryngée, l'évaluation perceptuelle (G.R.B.A.S.)^[122], l'évaluation du patient (V.H.I.)^[19] et le D.S.I.
[355]

CHAPITRE >

Discussion générale

La revue de la littérature nous a permis de détailler les outils mis à notre disposition dans le cadre de l'étude de la voix. Outils tels que la vidéostroboscopie, les échelles d'auto-évaluation, les échelles perceptuelles, les mesures objectives acoustiques (fréquence fondamentale, intensité, mesures de perturbation de la fréquence et de l'intensité, ...), les mesures objectives aérodynamiques (temps maximum phonatoire, capacité vitale, quotient phonatoire, le flux...), ainsi que les index vocaux. Chacun de ces outils est en perpétuel remaniement et bénéficie des apports technologiques.

L'évaluation stroboscopique dépend de l'œil averti de l'O.R.L., de son expérience et de son bagage professionnel. Plusieurs protocoles de base (B.S.G.V.D., S.E.R.F., S.A.V. [annexes 3, 4 et 5]) sont à l'étude pour standardiser les observations. Pour rappel, les paramètres observés sont l'amplitude, l'ondulation de la muqueuse, le degré de fermeture et de symétrie glottique, la régularité de la vibration, la nature et le comportement vibratoire de la lésion. Jusqu'à présent, la fiabilité de ces différents paramètres est moyenne.

De plus, la stroboscopie repose sur un effet d'optique. La cinématographie à haute vitesse captant jusqu'à 10.000 images par seconde est donc plus proche de la réalité. Cependant, elle reste encore confinée à quelques laboratoires vu son coût et les difficultés techniques qu'elle impose.

Toutefois, la plupart des auteurs (Dejonckere et al (1998)^[64], Ramos et al (1993)^[244], Teitler (1995)^[317], Poburka (1999)^[236]) précisent qu'un entraînement de base minimum serait susceptible d'améliorer la détermination de critères fiables et la précision des mesures.

La vidéostroboscopie reste donc actuellement l'outil de prédilection pour l'observation du larynx.

Les **échelles d'auto-évaluation** permettent au patient de décrire l'impact de leur dysphonie dans la vie courante et permet au thérapeute de mieux cibler les objectifs de la thérapie par rapport au desiderata du patient. Les premières échelles comportaient peu de critères d'évaluation et ne répondaient pas nécessairement aux situations concrètes de la vie du sujet. Le V.H.I. (Voice Handicap Index, Jacobson et al (1997)^[152] [annexe 5]) a l'avantage de quantifier l'aspect "invalidant" d'un éventail

de troubles vocaux sur la qualité de la vie. Sa fiabilité est haute (0.95). Il se corrèle aisément avec des échelles psychologiques, en particulier le SF-36, qui évalue la répercussion de la pathologie notamment sur la fonction sociale, la santé mentale et les émotions (Benninger et al (1998)^[19]). L'utilisation du V.H.I. se répand et satisfait les cliniciens des laboratoires de recherches. Il enrichit l'anamnèse du patient qui cible rapidement les difficultés rencontrées dans sa vie quotidienne. Le V.H.I. s'adapte adéquatement aux troubles vocaux les plus fréquemment rencontrés lors des consultations phoniatriques. Par contre, il traduit moins bien des perturbations plus discrètes telles que les dysodies d'où la nécessité d'adapter le questionnaire aux besoins des professionnels de l'art de la parole et du chant. Actuellement, la seule évaluation pour les chanteurs se résume en une anamnèse approfondie mise au point par Sataloff sous forme d'un questionnaire fouillé [in 267, p.877-884].

L'intégration du V.H.I. au bilan vocal a pour avantage d'inclure le patient dans l'examen en révélant au thérapeute la perception qu'a le sujet du degré de sévérité de sa dysphonie. Le V.H.I. présente de plus l'avantage d'avoir été éprouvé sur une large population.

Parmi les **échelles perceptuelles**, l'échelle G.R.B.A.S. est sans nul doute la plus appréciée par la plupart des auteurs. Sa facilité d'emploi en clinique journalière, la description cohérente de ses paramètres facilitent son utilisation pour tout professionnel (De Bodt et al (1996)^[51]). Les auteurs (Abe et al (1986)^[1], Leinonen et al (1997)^[181], De Bodt (1997)^[52], Kreiman et al (1993)^[171], Olson et al (1998)^[220], Millet et al (1998)^[201], Morsomme et al (1998)^[208], Carding et al (1999)^[35]) sont unanimes quant au degré de fiabilité du grade. Ils sont plus nuancés pour les paramètres de raucité et de souffle et trouvent des concordances plus faibles pour les paramètres de sensation de faiblesse et de sensation de forçage. Les échantillons sonores utilisés, l'expérience, la sensibilité de l'auditeur influencent l'évaluation perceptuelle. Aussi, certains auteurs, notamment Giovanni (1998)^[98], insistent pour que les stratégies d'écoute soient approfondies. La plupart du temps, le matériau sonore retenu est un échantillon spontané de parole qui permet de rendre compte de la dynamique vocale en condition naturelle (Hammarberg et al 1980^[108], (1984)^[109], (1986)^[110], (1992)^[111], de Krom et al (1994)^[54], Revis et al (1999)^[256], Giovanni (1998)^[93]).

L'échelle G.R.B.A.S. s'adapte bien aux différents troubles vocaux tant pour les adultes que pour les enfants. Cependant, pour les seuls troubles de la voix chantée, elle

s'avère peu efficace. La voix parlée de certains patients dysodiques est notée comme "état vocal normal" sur l'échelle G.R.B.A.S., alors qu'en chant, du souffle, de la raucité et de l'irrégularité vibratoire peuvent être perçus. Entre musiciens, chanteurs, phoniâtres, psycho-acousticiens, phonéticiens,...un même vocabulaire musical ne recouvre pas toujours une même réalité perceptuelle. Sur le plan perceptuel, aucune échelle à l'instar de la G.R.B.A.S. n'a encore été proposée pour évaluer la voix chantée. La tâche est aussi difficile que pour la dysphonie. Notre contribution a porté sur une caractéristique de la voix chantée qui est le vibrato (étude 1). Pour cette étude, tous les juges experts et les sujets de niveaux différents appartiennent au monde des chanteurs.

Pour assurer le fondements des résultats obtenus, nous avons établi que :

- Un haut niveau d'expertise ne renforce pas la concordance dans l'évaluation du vibrato. Les juges ont une réalité perceptuelle propre de chacun des paramètres et ne se rencontrent pas en un consensus commun. Pourtant les paramètres d'évaluation choisis tels que "harmonieux", "non-gênant", "régulier", "stable", "agréable" ont été décrits et discutés entre professionnels qui ont donné l'illusion d'une compréhension commune, mais la réalité perceptuelle qu'ils recouvrent a prouvé que cette unanimité n'est qu'illusoire.
- La reproductibilité du choix catégoriel (vibrato, son blanc, tremolo, chevrottement, autre) est meilleure pour les catégories vibrato et son blanc. Les juges confondent tremolo et chevrottement. La catégorie "autre" a été judicieusement utilisée ; les juges y ont classé les émissions instables incluant plusieurs types de trémulation.
- Sur le plan de la reproductibilité de l'échelle-test portant sur les paramètres ("gênant/non-gênant", "harmonieux/dysharmonieux" ; "agréable/désagréable", "régulier/irrégulier" ; "stable/instable" ; "timbré/détimbré"), le paramètre "timbré/détimbré" récolte un taux élevé de concordance pour 8 juges sur 10. Son aspect global le distingue des autres paramètres. Les paramètres "non-gênant/gênant", "harmonieux/dysharmonieux", "agréable/désagréable" constituent le facteur esthétique. Les paramètres "régulier/irrégulier", "stable/ instable" constituent un facteur d'ordre physico-acoustique. Le timbre, lui, fait partie du vocabulaire de tout chanteur qu'il soit débutant ou professionnel accompli.

- La complexité de la tâche (morceau imposé versus voyelle isolée) influence l'évaluation perceptuelle du vibrato. Ainsi que pour l'évaluation perceptuelle des dysphonies, le morceau imposé, tout comme le discours spontané, permet au jury de mieux appréhender la dynamique vocale d'ensemble de l'extrait, ceci engendre une bonne concordance entre les juges et ce, particulièrement pour le paramètre "timbré".

Bien que controversée, de par sa grande variabilité intra et inter-juges, l'évaluation perceptuelle reste non seulement un outil prisé par le thérapeute, mais plus encore un référent pour valider l'examen objectif. Même si, l'oreille semble peu fiable, particulièrement lorsque la tâche se complexifie notamment dans l'analyse d'une caractéristique vocale aussi pointue que le vibrato, tout comme Hammarberg (1980)^[108], nous sommes convaincus que bien entraînée, elle restera un moyen important d'évaluation vocale.

L'échelle G.R.B.A.S, largement détaillée et employée dans nos travaux, s'avère sur le plan clinique l'outil de prédilection. Les facteurs tels que le grade, la raucité et le souffle sont valides et corrélés aux paramètres objectifs de variabilité de la fréquence et de l'intensité, ainsi qu'au rapport signal/bruit.

La nécessité d'un entraînement spécifique à l'écoute, l'élaboration d'un consensus, et d'une définition claire de chaque paramètre basé sur des recherches scientifiques ne peuvent qu'affiner le jugement.

Il reste à tester et à améliorer l'échelle proposée dans notre étude 1 pour la voix chantée. L'exigence vocale et le rendement auxquels les chanteurs sont soumis requièrent une approche différente de l'expertise perceptuelle. Aussi tout comme pour l'échelle G.R.B.A.S., un entraînement spécifique est nécessaire pour obtenir un meilleur consensus.

Le développement des logiciels a permis un accès rapide à de multiples données acoustiques et aérodynamiques.

La voix normale a été observée à l'aide du SMLT (spectre moyen à long terme) notamment pour l'étude de l'influence du sexe, du stress, des langues étrangères, pour l'identification des caractéristiques spectrales des voyelles, de la mise en

vibration ou "voice onset", pour l'évolution de la voix en fonction de l'âge (Harmegnies (1988)^[112], (1992)^[113]).

L'analyse spectrale a été également appliquée aux voix pathologiques. Les travaux sur la voix des chanteurs et des locuteurs professionnels ont permis l'identification et la description du formant du chanteur, la précision des registres, la corrélation avec le degré d'aperture buccale, l'influence des mouvements du voile du palais, de la langue, la configuration du vibrato et l'impact de différentes techniques vocales sur l'analyse spectrale (Morris et al (1995)^[203], Miller (1986)^[200], Miller et al (1990)^[198], (1993)^[199], Novak et al (1995)^[217], Nawka et al (1997)^[212], Omori et al (1996)^[221], Siegwart (1995)^[283], Sundberg (1991)^[308], Griffin et al (1995)^[104], Harmegnies (1988)^[112], (1992)^[113], Blaylock (1999)^[23], Doskov et al (1995)^[68], Ekholm et al (1998)^[71]). Les courbes spectrales recueillies en période de pré et post rééducation ont permis d'objectiver les effets thérapeutiques. Quant à la fiabilité du SMLT, Harmegnies a pu montrer que la stabilité intra-locuteur des sujets est bonne pour une période de trois jours à trois semaines. Passé ce délai, l'on observe une variabilité graduelle. Les travaux portant sur l'analyse spectrale à long terme ont donc démontré son intérêt.

L'analyseur dynamique de signal Hewlett Packard^[117] est l'appareil que nous avons utilisé dans le cadre de notre étude (étude 2) sur l'analyse spectrale appliquée à la voix chantée (Morsomme et al (1993)^[205]) et appliquée à une population de sujets âgés (Morsomme et al (1997)^[206]). L'analyse spectrale a mis en évidence le formant du chanteur et le vibrato. Le formant est confirmé par l'enregistrement de la voix placée dans les résonateurs ; il s'observe particulièrement sur un fondamental grave ; le chanteur l'acquiert par la technique et l'expérience.

L'écoute doit précéder l'observation du vibrato, sinon sa description graphique pourrait sembler dénoncer un état pathologique. En effet la configuration des harmoniques en dents de scie peut faire penser à de la bitonalité, telle qu'on la rencontre dans les voix dysphoniques.

Selon le classement proposé par Remacle et al (1989)^[248], avec l'âge, les hommes ont une meilleure qualité vocale que les femmes âgées. L'analyse spectrale met en évidence une diminution de la fréquence fondamentale chez la femme et une augmentation de celle-ci chez l'homme. Les hommes âgés ont un moins bon

rendement vocal mais une meilleure qualité du vibrateur vocal que les femmes âgées (étude 3).

Des mesures aérodynamiques "manuelles" (comme décrites au chapitre 4 paragraphe 4.1.) comportant le temps maximum phonatoire, la capacité vocale et le quotient phonatoire ont été réalisées sur 40 sujets âgés de 70 à 90 ans. Nous avons établi une comparaison avec une population de sujets jeunes sur base des valeurs rapportées par Inglis et al (1985)^[144] et Taylor et al (1985)^[317] (étude 3).

La capacité vitale et le temps maximum phonatoire diffèrent selon le sexe et l'âge, tandis que le quotient phonatoire ne révèle pas de différence significative entre hommes et femmes âgés. Nous constatons une nette diminution des performances des sujets âgés pour la capacité vitale et la tenue du [a] par rapport aux sujets jeunes. Bien que la capacité physique des femmes âgées soit réduite, leur rendement vocal reste bon par rapport aux femmes jeunes. Les hommes quant à eux montrent une dégradation du rendement vocal de 25% par rapport à celui des jeunes. Ces résultats corroborent ceux de la littérature (Morsomme et al (1997)^[206]).

Les logiciels actuels proposent une quantité de données régies par des algorithmes différents d'un appareil à l'autre. Nous ne disposons actuellement d'aucun standard. Nous avons successivement travaillé avec les stations E.V.A. 1^[293] et E.V.A. 2^[296] (Évaluation Vocale Assistée) qui fournissent des mesures de fréquence fondamentale, d'intensité, de jitter et de shimmer, ainsi que des mesures aérodynamiques, de débit d'air buccal et de fuite glottique.

Nous avons pu établir sur une population de 103 enfants âgés de 4 à 9 ans des valeurs de référence pour chacun des paramètres analysés par la station E.V.A.. Notre analyse s'est portée d'une part sur la partie jugée la plus stable de la voyelle [a] et sur le début de l'émission sonore. Une comparaison avec un groupe de choristes, nous a permis d'observer l'effet positif de l'entraînement de la voix chantée. Les chanteurs ont un fondamental systématiquement plus élevé que les non chanteurs et ce, de 9 HZ.

Le placement de la voix dans les résonateurs fait partie de l'éducation chantée du choriste dès le début de l'apprentissage. Dans le chant, les valeurs de consommation d'air et de débit d'air buccal sont plus élevées, nous avons pu montrer que

l'inexpérience du chant est mise en évidence par des mesures plus élevées de shimmer et d'intensité écart (étude 5).

La mesure de la pression sous-glottique estimée selon la méthode de Smitheran et Hixon (1981)^[286] s'intègre facilement à la clinique contrairement à la mesure de la pression sous-glottique par ponction au travers du ligament crico-trachéal. Celle-ci est plus précise mais invasive.

Différentes études (Löfquist et al (1982)^[186], Hertegard et al (1995)^[115]) ont montré que les valeurs de pression intra-orale, de pression sous-glottique et de pression sous-glottique estimée sont hautement corrélées. Les discussions se sont davantage orientées sur le degré de niveau de pression sonore émise par le sujet.

Pour les voix pathologiques, l'interprétation des valeurs doit être prudente, notamment pour les voix dont la qualité perceptuelle est soufflée^[98].

La station E.V.A. 2^[296] propose la mesure de la pression sous-glottique estimée. Aussi, avons-nous réalisé nos premières mesures de pression sur un échantillon de 88 sujets adultes âgés de 20 à 50 ans en vue d'établir des valeurs de référence (étude 7). Nos résultats hommes/femmes (7.46 hpa/ 7.70 hpa) sont similaires à ceux de la littérature (Stathopoulos et al (1985)^[298], Netzell (1969)^[214], Giovanni (1998)^[98]) ; afin d'éprouver nos valeurs, nous les avons comparées à celles d'un groupe de 15 sujets féminins souffrant de nodules ou d'un œdème de Reinke. Le groupe normal affichait une valeur moyenne de PSGE de 7.8 hpa, le groupe nodule de 10.71 hpa et le groupe œdème de Reinke de 9.7 hpa. Une pression sous-glottique estimée anormalement élevée chez les sujets dysphoniques est cohérente et corrobore les résultats obtenus par Zeitels (1997)^[361] et Giovanni (1998)^[98]. La mesure de la PSGE permet de différencier les sujets dysphoniques des sujets témoins.

Nous pensons que la mesure aérodynamique de la pression sous-glottique estimée va se perfectionner et constituer un critère discriminant majeur. Des études ont confirmé la pertinence de l'estimation de la pression sous-glottique par la pression intra-orale. Celle-ci rendrait compte de la sensation de forçage. Un consensus sur le placement de la sonde buccale validerait la mesure et permettrait d'établir des valeurs de référence.

Dans la littérature, les corrélations entre mesures subjectives et mesures objectives les plus fréquemment répandues concernent surtout les mesures de variabilité de la fréquence, et de l'intensité, soit respectivement le jitter et le shimmer ; ces corrélations se trouvent aussi entre le rapport signal/bruit et les mesures perceptuelles telles que les paramètres de raucité, souffle, enrouement et grade, entendu comme l'aspect général de la dysphonie. Les résultats sont hautement variables en fonction des auteurs (Deal et al (1978)^[50] ; Kojima et al (1980)^[167] ; Yumoto et al (1984)^[358] ; Wolfe et al (1995)^[351], (1997)^[352, 354] ; Eskenazi et al (1990)^[77] ; Dejonckere et al (1993)^[61] ; de Krom (1995)^[55] ; Millet et al (1998)^[201]), des outils logistiques, de leurs algorithmes de calcul, du degré de fiabilité du jury d'écoute, des tâches des sujets et de leur nombre.

Nous avons comparé 6 paramètres vocaux objectifs mesurés par le M.D.V.P. (Fftr, Fatr, Ftri, Atri, Jitt et Shim) avec une échelle perceptuelle construite pour l'évaluation de la qualité du vibrato ("gênant/non-gênant", "harmonieux/dysharmonieux" ; "agréable/désagréable", "régulier/irrégulier" ; "mature/embryonnaire" ; "précis, imprécis" "stable/ instable" ; "timbré/ détimbré") (étude 4) (Morsomme et al (1999)^[207]).

Cent vingt échantillons ont été appréciés par 3 juges. Un seul juge présente une concordance satisfaisante. Ses réponses sont significativement corrélées avec les paramètres de Ftri, Jitt et Shim. On peut conclure que l'oreille avec le talent, l'entraînement et l'expérience est sensible à certains paramètres vocaux, cependant il semble que ce soit une exception.

Nous avons comparé les mesures vocales objectives obtenues par la station E.V.A. 1^[293] avec les paramètres de l'échelle G.I.R.B.A.S.. Huit juges experts ont évalués 40 échantillons vocaux provenant de 28 sujets atteints de paralysie laryngée unilatérale et de 12 sujets formant le groupe témoin (étude 6). Les paramètres objectifs exprimant l'apériodicité du signal à savoir le coefficient de variabilité du fondamental (Fo coef) et de l'intensité (INT coef) ainsi que le jitter (jitt) sont liés aux dimensions de grade, de souffle et de sensation de faiblesse. La réalité perceptive d'une dysphonie occasionnée par une paralysie laryngée relève davantage de ces trois paramètres (G, B, A) que de ceux de raucité (R) et d'instabilité (I).

L'avantage de l'examen objectif est que contrairement au perceptuel, il n'est pas influencé par la tâche, l'ordre de la tâche et la disponibilité d'écoute du juge. Cet

examen doit rendre compte d'une signification physiopathologique la plus claire possible. Cependant, les mesures de variabilité telles que le jitter et le shimmer sont limitées. La première difficulté est dépendante de la méthode d'extraction de la fréquence fondamentale qui diffère d'un appareil à l'autre, rendant la comparaison utopique et inutile. La seconde remet en cause la pertinence de ces mesures pour des signaux irréguliers et ce particulièrement pour le jitter. Au-delà d'un certain pourcentage de variabilité, la mesure n'est plus fiable.

Les mesures objectives doivent donc être interprétées dans leur contexte, et dépendent directement de la fiabilité et de la validité de l'instrument employé. De plus concernant la tâche, l'échantillon idéal à analyser serait la parole spontanée. Toutefois, la voix en co-articulation de par les mises en vibration et les arrêts vibratoires nécessite une analyse subtile et fine, auquel aucun programme n'a encore réellement répondu.

CHAPITRE 8

Conclusion

A l'issue de ce travail, nous pensons que le bilan vocal actuel, à l'instar de De Bodt (1997)^[52] doit comporter :

- une évaluation stroboscopique,
- une évaluation perceptuelle,
- une auto évaluation,
- des mesures objectives,
- le calcul du D.S.I.^[355].

Sur le plan de la voix parlée, l'examen du patient, selon ce canevas reste dans des délais raisonnables et ne comporte qu'une seule investigation invasive : la stroboscopie. L'avis du patient et du thérapeute y sont pris en compte et un index tel que le D.S.I. qui reste encore à définir, apportera une vue plus globale de la sévérité du trouble traité.

Sur le plan de la voix chantée, nous proposons un canevas similaire à celui de la voix parlée. Toutefois, une échelle d'auto-évaluation adéquate devrait être mise au point pour cibler les difficultés rencontrées par les artistes chanteurs, comédiens et autres locuteurs chevronnés dans l'exercice de leur pratique professionnelle. De même l'échelle perceptuelle gagnerait à être plus spécifique et adaptée aux exigences du métier. Nous songeons particulièrement aux caractéristiques du vibrato et du formant du chanteur qui devraient s'inclure dans l'évaluation vocale, ceci tant sur le plan objectif par l'analyse spectrale et des mesures objectives appropriées, que sur le plan subjectif dans l'appréciation d'un facteur général (le timbre) et de facteurs esthétique et physico-acoustique. Cette thèse a approfondi le domaine de la voix chantée jusqu'ici parent pauvre des recherches et de la phoniatrie, nos résultats constituent une étape. Ils nous permettent de proposer une série d'autres investigations qui affinaient les mesures de vibrato notamment en ce qui concerne sa cadence et son amplitude et qui enrichiraient l'étude des paramètres de notre échelle perceptuelle.

Nos résultats concernant les mesures objectives et les corrélations avec les mesures perceptuelles ont mis en évidence l'intérêt de recourir aux logiciels pour standardiser les différents paramètres qui analysent la voix. Etablir des tables de valeurs de référence pour chaque logiciel employé offre aux cliniciens la possibilité de comparer les résultats de ses patients à un standard de référence qui devrait par exemple tenir compte des méthodes de calcul d'extraction de la fréquence fondamentale.

Par ailleurs, les logiciels aussi perfectionnés et fiables soient-ils ne resteront que des outils au service du praticien qui affinera ses qualités d'écoute de la voix, et d'analyse du geste vocal en vue d'établir un diagnostic et un programme thérapeutique toujours plus efficaces.

Bibliographie

A

1. Abe H, Yonekawa H, Ohta F, Imaizumi S. Reproducibility of hoarse voice psychoacoustic evaluation. *Jpn. J. Logoped. Phoniat.* 1986;27:168-177.
2. Abitbol J, Abitbol P, Abitbol B, Sex hormones and the female voice. *J Voice* 1999; 3: 424-446.
3. Alipour F, Titze I. Active and passive characteristics of the canine cricothyroid muscles; 1999, 13/1: 1-10.
4. Anders L, Hollien H, Hurme P, Sonninen A, Wendler J. Perception of hoarseness by several classes of listeners. *Folia Phoniatr* 1988;40: 91-100.
5. Andreassen ML, Smith BE, Guyette TW. Pressure-flow measurements for selected oral and nasal sound segments produced by normal adults. *Cleft Palate Craniofac J.* 1992; 29/1:1-9.
6. Andrews ML, Schmidt Ch P. Gender presentation: perceptual and acoustical analyses of voice. *J Voice* 1997; 11/3: 307-313.
7. Arkebauer HJ, Hixon TJ, Hardy JC. Peak intraoral air pressure during speech. *J Speech Hear Res* 1983;26: 202-208.
8. Aronson AE. *Clinical voice disorders* (2nd ed). Thieme. New York .1985; 265 p.
9. Aronson AE. Importance of the Psychosocial interview in the diagnosis and treatment of “functional” voice disorders. *J Voice* 1990;4: 287-289.
10. Atkinson JE. Correlation analysis of the physiological factors controlling fundamental voice frequency. *J Acoust Soc Am* 1978;63: 211-222.
11. Avery JD, Liss JM. A perceptual and acoustic analysis of less-masculine-sounding speech. Paper presented at the Annual Convention of the American Speech Language and Hearing association, Anaheim, CA, Nov 1993.
12. Awan SM. Phonetographic profiles and Fo-SPL characteristics of untrained versus trained vocal groups. *J Voice* 1991;5: 41-50.

B

13. Baer T. Reflex activation of laryngeal muscles by sudden induced subglottal pressure changes. *J Acoust Soc Am* 1979;65: 1271-1275.

14. Baken RJ, Orkillo RF. Phonatory response to step-function changes in supraglottal pressure. In: Baer T, Sasaki C, Harris K, eds. *Laryngeal function in phonation and respiration*. Boston: College-Hill Press, 1987: 273-290.
15. Baken RJ. Clinical measurement of speech and voice. Boston : College Hill, 1987 : 171-182.
16. Baken JR. An overview of laryngeal function for voice production. In: Sataloff RT , *Professional Voice. The science and art of clinical care* (2nd ed) Singular Publishing Group, Inc London 1997 p147-166.
17. Barth V. Die laryngoskopie als Möglichkeit der funktionsdiagnostik von stimmstörungen und stimmritzenprozessen. Richard Wolf GmbH Knittlingen 1982.
18. Bassich CJ, Ludlow C. The use of perceptual methods by new clinicians for assessing voice quality. *J Speech Hear Res* 1986;51: 125-133.
19. Benninger MS, Ahuja AS, Gardner G, Grywalski C. Assessing outcomes for dysphonic patients. *J Voice* 1998 ;12 : 540-550.
20. Bernthal JE, Beukelman DR. Intraoral air pressure during the production of /p/ and /b/ by children, youths and adults. *J Speech Hear Res* 1978;21: 361-367.
21. Bielamowicz S, Kreiman J, Gerratt BR, Dauer MS, Berke GS. Comparison of voice analysis systems for perturbation measurement. *J Speech Hear Res* 1996 ;39 : 126-134.
22. Biondi S, Zappala M, Amato G, Consoli F. La voce senile. *Acta Otorhinol Ital* 1992; 12: 69-79.
23. Blaylock TR. Effects of systematized vocal warm-up on voices with disorders of various etiologies. *J Voice* 1999 ;13 : 43-50.
24. Bless D, Hirano M, Feder RJ. Videostroboscopic evaluation of the larynx. *Ear Nose Throat J* 1987;66: 289-296.
25. Böhme G, Stuchlik G. Voice profiles and standard voice profile of untrained children. *J Voice* 1995 ;9 : 304-307.
26. Boone D, McFarlane S. *The voice and voice therapy*, (5th ed). Englewood Cliffs, NJ: Prentice Hall, 1994.
27. Boone DR. The singing/acting voice in the mature adult. *J Voice* 1997 ;11 : 161-164.
28. Bouchayer M, Cornut G. Microsurgical treatment of benign vocal fold lesions: indications, technique, results. *Folia Phoniatr* 1992; 44: 155-184.

29. Bough ID, Heuer JR, Sataloff RT, Hills JR, Cater JR. Intrasubject variability of objective voice measures. *J Voice* 1996 ;10 : 166-174.
30. Brown WS. Supraglottal air pressure variations associated with consonant productions by children. *Curr Issues Ling Theory* 1979;9: 935-944.
31. Brown WS, Morris RJ, Michel J. Vocal jitter and fundamental frequency characteristics in aged female professional singers. *J Voice* 1990 ;4 : 135-141.
32. Brown WS, Morris RJ, Hollien H, Howell E. Speaking fundamental frequency characteristics as a function of age and professional singing. *J Voice* 1991 ;5 : 310-315.
33. Brown WS, Morris RJ, Hicks DM, Howell E. Phonation profiles of female professional singers and non-singers. *J Voice* 1993 ;7 : 219-226.

C

34. Caitriona MC, Mc Hugh-Meunier, Klaus R, Schener, Lehmann W, Scherer U, Coping strategies, personality and voice quality in patients with vocal folds nodules and polyps. *J Voice* 1997;4: 452-461.
35. Carding PN, Horsley IA, Doherty GJ. A study of the effectiveness of voice therapy in the treatment of 45 patients with non-organic dysphonia. *J Voice* 1999 ;13 : 72-104.
36. Case J. Clinical management of voice disorders, (2nd ed). Austin, TX : Pro-ed, 1991.
37. Chodzko-Zajko WJ, Ringel RL. Physiological aspects of aging. *J Voice* 1987;1: 18-26.
38. Cleveland TF, A clearer view of singing voice production: 25 years of progress. *J Voice*, 1994; 8: 18-23.
39. Cleveland TF, Tone RE, Sundberg J, Iwarsson J. Estimated subglottal pressure in six professional country singers. *J Voice* 1997 ;11 : 403-409.
40. Colcord RD, Adams MR. Voicing duration and vocal speech changes associated with stuttering reduction during singing. *J Speech Hear Res* 1979 ;22 : 468-479.
41. Coleman RF, Mabis JH, Hinson JK. Fundamental frequency-sound pressure level profiles of adult male and female voices. *J Speech Hear Res* 1977 ;20 : 197-204.

42. Coleman RF, Hicks DM, singer's compensation for varying loudness levels of musical accompaniment. In Lawrence V, ed. transcripts of the seventh symposium: care of the professional voice, part I: the scientific papers. New York: The Voice Foundation, 1978: 80-91.
43. Coleman RF. Sources of variation in phonetograms. J Voice 1993;7: 1-14.
44. Colton R, Casper J. Understanding voice problems: a physiological perspective for diagnosis and treatment. Baltimore, MD: Williams and Wilkins, 1990.
45. Cooper M. The tired speaking voice and the negative effect in the singing voice. NATS Bull 1982;39: 11-14.
46. Cornut G. Etude clinique de la voix dans le cadre du bilan phoniatrique. Oto-Rhino-Laryngol., 1991, 20753 A10: pp.8.
47. Costamagna D. La vidéo-laryngostroboscopie: techniques et recueil des données cliniques. Bull Audio-phonol. Ann Sc Univ, Franche-Comté 1990;6: 491-546.

D

48. Dalleas B, Dalleas A. Qualités acoustiques de la voix. Rev Laryngol 1987;108: 373-377.
49. D'Alessandro C, Castellengo M. Etude, par la synthèse, de la perception du vibrato vocal dans les transitions de notes. Bulletin d'audiophonologie. Ann. Sc. Univ. Franche-Comté 1991; 7, 5 & 6: 551-564.
50. Deal RE, Emanuel FW. Some waveform and spectral features of vowel roughness. J Speech Hear Res 1978;21: 250-264.
51. De Bodt MS, Van de Heyning Ph, Wuyts FI, Lambrechts L. The perceptual evaluation of voice disorders. Acta Otot-Rhino-Laryngologica belg 1996; 50: 283-291.
52. De Bodt MS, A framework for voice assessment. The relation between subjective and objective parameters in the judgment of normal and pathological voice. Doctoral Dissertation. 1997 University of Antwerp.
53. De Bodt MS, Wuyts FL, Van de Heyning PH, Croux C. Test-retest study of the GRBAS scale : influence of experience and professional background on perceptual rating of voice quality. J Voice 1997 ;11 : 74-80.

54. de Krom G. Consistency and reliability of voice quality ratings for different types of speech fragments. *J Speech Hear Res* 1994;37: 985-1000.
55. de Krom G. Some spectral correlates of pathological breathy and rough voice quality for different types of vowel fragment. *J Speech Hear Res* 1995 ;38 : 794-811.
56. Decoster W, Debruyne F. Changes in spectral measures and voice-onset time with age : a cross-sectional and a longitudinal study. *Folia Phoniatr Logop* 1997 ;49 : 269-280.
57. Dejonckere Ph. Le concept oscillo-impédantiel de la vibration laryngée. *Rev Laryngol* 1985; 106: 275-282.
58. Dejonckere PH, Dejong-Estienne Fr. Techniques de base d'évaluation de la voix, Cabay, 1985.
59. Dejonckere Ph, Villarosa D. Analyse spectrale moyennée de la voix. Comparaison de voix normales et de voix altérées par différentes catégories de pathologies laryngées. *Acta Otorhinolaryngol. Belg* 1986; 40 : 426-435.
60. Dejonckere PH. Bruit de turbulence et apériodicité dans la voix pathologique. Une approche multifactorielle. *Rev Laryngol* 1990 ;111 : 535-537.
61. Dejonckere P, Obbens C, De Moor GM, Wienecke GH. Perceptual evaluation of dysphonia: reliability and relevance. *Folia Phoniatr* 1993;45: 76-83.
62. Dejonckere PH. Een multicentrisch onderzoek over de perceptieve stembeoordeling betrouwbaarheid en klinische relevantie. *Proceedings Dag van de Foniatrie* 13/5/95, Utrecht, 1995: 2-4.
63. Dejonckere PH, Hirano M, Sundberg J. *Vibrato*, San Diego : California. Singular Publishing Group, 1995.
64. Dejonckere PH, Crevier L, Elbaz E, Marraco M, Millet B, Remacle M, Woisard V. Quantitative rating of video-laryngostroboscopy : a reliability study. *Rev Laryngol Otol Rhinol* 1998 ;119 : 259-260.
65. Deem JF, Manning WH, Knack JV, Matesich JS The automatic extraction of pitch perturbation using microcomputers: some methodological considerations. *J Speech Hear Res* 1989; 32: 689-697.
66. Dobinson CH, Kendrick AH. Normal values and predictive equations for aerodynamic function in British Caucasian subjects. *Folia Phoniatr*, 1993;45: 14-24.

67. Doherty E, Shipp T. Tape recorder effects on jitter and shimmer extractions. *J Speech Hear Res*, 1988;31: 485-490.
68. Doskov D, Ivanov T, Boyanov B. Comparative analysis of singer's high formant in different type of singing voices. *Folia Phoniatr Logop* 1995 ;47 : 291-5.
69. Drew R, Sapir S. Average speaking fundamental frequency in soprano singers with and without symptoms of vocal attrition. *J Voice* 1995 ;9 : 134-41.

E

70. Egger J, Friedl W, Friedrich G. Functional dysphonia: personality and coping behavior in stressful situations. In: Zapotoczky HG, Wenzel T, The scientific dialogue: from basic research to clinical intervention. Amsterdam: Swets and Zeitlinger, 1990: 201-207.
71. Ekholm E, Papagiannis GC, Chagnon FP. Relating objective measurements to expert evaluation of voice quality in western classical singing : critical perceptual parameters. *J Voice* 1998 ;12 : 182-196.
72. Ekman P, Firesen WV, Scherer KR. Body movement and voice pitch in deceptive interaction. *Semiotica* 1976;16: 23-27.
73. Ekstrom ER, Control of singing intensity as related to singer experience. *The NATS Bulletin* 1960;17: 8-12.
74. Elias ME, Sataloff RT, Rosen DC, Heuer RJ, Spiegel JR. Normal stroboscoped laryngoscopy : variability in healthy singers. *J Voice* 1997 ;11 : 104-107.
75. Elliot N, Sundberg J, Gramming P. What happen during vocal warm-up ? *J Voice* 1995 ;9 : 37-44.
76. Endres W, Bambach W, Flösser G. Voice spectrograms as a function of age, voice disguise and voice imitation. *J Acoust Soc Am* 1971;49: 1842-1848.
77. Eskenazi L, Childers DG, Hicks DM. Acoustic correlates of vocal quality. *J Speech Hear Res* 1990;33: 298-306.

F

78. Fant G. On the predictability of formant levels and spectrum envelopes from formant frequencies. In: Halle M, ed. For Roman Jakobson. S'Gravenhage: Mouton & Co., 1956: 109-120.
79. Fant G. Speech Sounds and features, Cambridge, London, Mit Press 1973.
80. Ferrand CT, Bloom RL. Gender differences in children's intonational patterns. J Voice 1996 ;10 : 284-291.
81. Fitch JL, Holbrook A. Modal vocal fundamental frequency of young adults. Archives of Otolaryngology 1970; 92: 379-382.
82. Foot EL. The effect of the intensity of auditory feedback on the loudness-intensity, quality-wave form and intelligibility of the singer's voice [Dissertation]. Lawrence, Kansas: university of Kansas, 1965.
83. Frank F, Donner F. Stimmefeldmessungen unter besonder Berücksichtigung der Tonhaldedauer bei Durchschnittsstimmen und Kunstgesangsstimmen; in Spitzer L (ed): Bericht über das 4. Gesangspädagogische Symposion, Bad Ischl, März 1983, pp 65-75.
84. Freidl W, Egger J, Friedrich G. Persönlichkeit und Stress-verbeitung bei funktionellen Dysphonikern. PsychotherPsychosom Med Psychol 1989; 39: 300-305.

G

85. Gauffin J, Sundberg J. Spectral correlates of glottal voice source wave form characteristics. J Speech Hear Res 1980;32: 556-565.
86. Gauffin J, Sundberg J. Data on the glottal voice source behavior in vowel production. In: Gramming P (Ed.). The phonetogram: an experimental and clinical study. Malmö, Doctoral Dissertation, 1988.
87. Gelfer MP. Perceptuel attributes of voice: Development and use of rating scales. J Voice 1988; 2: 320-326.
88. Gelfer MP, Fendel DM. Comparisons of jitter, shimmer, and signal-to-noise ratio from directly digitized versus tape voice samples. J Voice 1995 ;9 : 378-382.
89. Gelfer MP, Young SR. Comparisons of intensity measures and their stability in male and female speakers. J Voice 1997 ;11 : 178-186.

90. Gelfer PM, Schofield KJ. Comparison of acoustic and perceptual measures of voice in male to female transsexuals perceived as female versus those perceived as male. *J Voice*, 2000; 14, 1: 22-23.
91. Gerratt BR, Kreiman J, Antonanzas-Barroso N, Berke GS. Comparing internal and external standards in voice quality judgements *J Speech Hear Disord* 1993; 36 : 14-20.
92. Giovanni A, Molines V, N'guyen N. Une méthode multiparamétrique d'évaluation vocale objective assistée par ordinateur. *Ann Otolaryngol Chir Cervicofac* 1992; 109: 200-206.
93. Giovanni A. Analyse objective des dysphonies avec l'appareillage E.V.A.: étude préliminaire à la définition d'un protocole normalisé. Mémoire pour le diplôme de phonétique expérimentale, dirigé par Di Cristo A, 1994-1995.
94. Giovanni A, Estublier N, Robert D, Teston B, Zanaret M, Cannoni M. Evaluation vocale objective des dysphonies par la mesure simultanée de paramètres acoustiques et aérodynamiques à l'aide de l'appareillage EVA. *Ann Otolaryngol Chir Cervicofac* 1995 ; 112 : 85-90.
95. Giovanni A, Robert D, Estublier N, Teston B, Zanaret M, Cannoni M. Objective evaluation of dysphonia: preliminary results of a device allowing simultaneous acoustic and aerodynamic measurements. *Folia Phoniatr Logop* 1996;48: 175-185.
96. Giovanni A, Robert D, Teston B, Guarella MB, Zanaret M. Etude préliminaire des paramètres acoustiques et aérodynamiques après laryngectomie frontale antérieure de Tucker. *Ann Otolaryngol Chir Cervicofac* 1996;113: 277-284.
97. Giovanni A, Robert D, Teston B, Guarella MD, Estublier N, Zanaret M. L'analyse objective des dysphonies avec l'appareillage EVA : état des lieux et perspectives. *Les Cahiers d'O.R.L.* 1996 ;21 : 143-155.
98. Giovanni A. Contribution à l'élaboration d'un protocole multiparamétrique d'évaluation objective des dysphonies. Dissertation Doctorale. Université de la Méditerranée 1998.
99. Glaze LE, Bless DM, Milenkovic P, Susser RD. Acoustic characteristics of children's voice. *J Voice* 1988;2: 312-319.
100. Gould WJ, Korovin GS. Laboratory advances for voice measurements. *J Voice* 1994 ;8 : 8-17.

101. Gramming P, Sundberg J. Spectrum factors relevant to phonetogram measurement. *J Acoust Soc Am* 1988;83: 2352-2360.
102. Greene MCL. The voice and its disorders. Philadelphia : Lippincott, 1972.
103. Gregg JW. The singing/acting mature adult-singing instruction perspective. *J Voice* 1997 ;11 : 165-170.
104. Griffin B, Woo P, Colton R, Casper J, Brewer D. Physiological characteristics of the supported singing voice. A preliminary study. 1995.
105. Gross M, Collo D. Funktionelle Ergebnisse nach Larynxteilresektion mit kranial gestielter Halsfasziendeckung. *Arch Oto-Rhino-Laryngol* 1983; suppl II: 201-202; personal communication about the lecture at the 54. Jahresversammlung der Deutschen Gesellschaft für Hals-Nasen-Ohren-Heilkunde, Kopf-und Hals-Chirurgie, Travemünde, 1983.

H

106. Hacki T. Automatisches Stimmfeldmesssystem mit integriertem Stimmbelastungstest, Hamburg, Homoth, 1993.
107. Hakes J, Doherty E TH, Shipp TH. Trillo rates exhibited by professional early musik singers. *J Voice* 1990;4: 305-308.
108. Hammarberg B, Fritzell B, Gauffin J, Sundberg J, Wedin L. Perceptual and acoustic correlates of abnormal voice qualities. *Acta Otolaryngol* 1980 ;90 : 441-451.
109. Hammarberg B, Fritzell B, Schiratzki H. Teflon injection in 16 patients with paralytic dysphonia: perceptual and acoustic evaluations. *J Speech Hear Disord* 1984;49: 72-82.
110. Hammarberg B, Fritzell B. Acoustic and perceptual analysis of vocal dysfunction. *J Phonet* 1986;14: 533-547.
111. Hammarberg B. Pathological voice qualities. Perceptual and acoustic characteristics of a set of Swedish "reference" voices. *Bull Audiophonol* 1992;VIII: 39-52.
112. Harmegnies B. Contribution à la caractérisation de la qualité vocale. Analyses plurielles de spectres moyens à long terme de la parole. Dissertation doctorale 1987-1988.

113. Harmegnies B. Les sources de variation du spectre à long terme de parole: revue de littérature. *Canadian Acoustics* 1992, 20: 9-35.
114. Heaver L. Psychiatric observations on the personality structure of patients with habitual dysphonia. *Logos* 1958;1: 21-26.
115. Hertegard S, Gauffin J, Lindestad PA. A comparison of subglottal and intra oral pressure measurements during phonation. *J. of voice*, 1995; 9, 2,: 149-155.
116. Heuillet-Martin G, Garson-Bavard H, Légré A. Une voix pour tous. La voix normale et comment l'optimiser (Tome1, 2^{ème} éd) Solal Editeur 1997: pp 204.
117. Hewlett-Packard model 3561 A, Application notes, Palo Alto, California.
118. Heylen LG, Wuyts FL, Mertens FW, Pattyn JE. Phonetography in voice diagnoses. *Acta O.R.L. Belg* 1996;50: 299-308.
119. Heylen LG, De klinische relevantie van het fonetogram. Een onderzoek bij kinderen en leerkrachten, Antwerpen, Doctoraal proefschrift U.I.A., 1997.
120. Higgins MB, Saxman JH, A comparison of selected phonatory behaviors of healthy aged and young adults, *J Speech and Hear research*, 1991; 34.
121. Hirano M, Koike Y, von Leden H. Maximum phonation time and air usage during phonation. *Folia Phoniatri*, 1968;20: 185-201.
122. Hirano M. Clinical examination of voice. New York, Springer Verlag, 1981: 81-84.
123. Hirano M, Kurita S, Nakashima T. Growth, development, and aging of human vocal folds. In Bless D, Abbs J, eds. *Vocal fold physiology: contemporary research and clinical issues*. San Diego, Calif: College-Hill Press, 1983: 22-43.
124. Hirano M, Hibi S, Teresawa R, Fujiu M. Relationship between aerodynamic, vibratory, acoustic and Psychoacoustic correlates of dysphonia. *J Phonet* 1986;14: 445-56.
125. Hirano M. Objective evaluation of the human voice clinical aspects. *Folia Phoniatri* 1989 ;41 : 69-144.
126. Hirano M. Clinical applications of voice tests. In: NIDCD, ed. *Assessment of speech and voice production*. NIDCD, Maryland, 1990: 196-203.
127. Hirano M, Bless D. *Videostroboscopic examination of the larynx*. Whurr Publishers, London, 1993.
128. Hirose H. High-speed digital imaging of vocal fold vibration. *Acta ORL (Stockh)*, 1998; 458: 151-153.

129. Hixon TJ, Klatt DH, Mead J. Influence of forced transglottal pressure changes on vocal fundamental frequency. *J Acoust Soc Am* 1971;49: 105(A).
130. Hollien H, Shipp T. Speaking fundamental frequency and chronological age in males. *J Speech Hear Res* 1972;15: 155-159.
131. Hollien, Doherty M, Doherty E. A method for analyzing vocal jitter in sustained phonation. *J Phonetics* 1973;1: 85-91.
132. Hollien H. On vocal registers. *J Phonet* 1974;2: 125-143.
133. Hollien H. Differentiating between speaking and singing vocal registers. In: Broecke M, Cohen A, eds. *Proceedings of the Tenth International Congress on Phonetic Sciences*. Dordrecht, Fons Publ., 1984: 417-422.
134. Hollien H. Report on vocal registers. In: Askenfelt A, Felicetti S, Jansson E, Sundberg J, eds. *Proceedings, Stockholm Music Acoustics Conference*. Royal Academy of Music, Stockholm, Sweden, 1985: 27-35.
135. Hollien H. Old Voices: what do we really know about them? *J. Voice* 1987; 1: 2-17.
136. Hollien H. That golden voice-Talent or training ? *J Voice* 1993 ;7 :195-205.
137. Holmberg E, Hillman R, Perkell J. Glottal airflow and transglottal air pressure measurements for male and female speakers in soft, normal, and loud voice. *J Acoust Soc Am* 1988;84: 511-529.
138. Horii Y. Fundamental frequency perturbation observed in sustained phonation. *J speech Hear Res*, 1979;22: 5-19.
139. Horii Y, Ryan W. Fundamental frequency characteristics and perceived age of adult speakers. *Folia Phoniatr* 1981 ;33 : 227-233.
140. Horii Y. Jitter and shimmer differences among sustained vowel phonations. *J speech Hear Res*, 1982;25: 12-14.
141. Hudson A, Holbrook A. Fundamental frequency characteristics of young Black adults: spontaneous speaking and oral reading, *Journal of Speech and Hearing Research*, 1982, 25: 25-28.
142. Husson R. *La voix chantée*. Paris. Gauthier-Villars, 1960.



143. Ingrisano D, Weismer G, Shuckers GH. Sex identification of preschool children's voices. *Folia Phoniatr* 1980;32: 61-69.

144. Inglis JM, In Johnson TS. Voice disorders: The measurement of clinical progress. Costello JM, (ed.) 1977. Speech disorders in adults. Recent advances. San Diego: College-Hill Press, 1985: 127-152.
145. Isshiki N. Regulatory mechanism of the pitch and volume of voice. *Practica Otol* 1959;52: 1065-1094.
146. Isshiki N. Vocal intensity and air flow rate. *Folia Phoniatr* 1964;5: 92-104.
147. Isshiki N, Yanagihara N, Morimoto M. Approach to the objective diagnosis of hoarseness. *Folia Phoniatr* 1966;18: 393-400.
148. Isshiki N, Okamura H, Tanabe M, Morimoto M. Differential diagnosis of hoarseness. *Folia Phoniatr* 1969;21: 9-12.
149. Isshiki N, Takeuchi Y. Factor analysis of hoarseness. *Stud Phonol* 1970;5: 37-44.
150. Iwata S. Aerodynamic aspects for phonation in normal and pathologic larynges. In: Fujimura O, ed Voice production mechanism and functions. New York, NY: raven Press, 1988: 423-433.
151. Izdebski K. Long-Time-Average-Spectra (S) applied to analysis of spastic dysphonia; in Lawrence V (ed): Transcripts 9th Symp. Care of the Professional Voice. New York, Voice Fondation, 1980, pp 89-144.

J

152. Jacobson BH, Johson A, Grywalski C, Silberglent A, Jacobson G, Benninger MS, Newman C. The voice handicap index (VHI): Development and validation. *American J of Speech Pathology* 1997; 6: 66-70.
153. Jafari M, Till JA, Truesdell LF, Law-Till CB. Time-shift, trial, and gender effects on vocal perturbation measures. *J Voice* 1993
154. Jiang J, O'Mara T, Conley D, Hanson D. Phonation threshold pressure measurements during phonation by airflow interruption. *Laryngosc* 1999 ;109 : 425-32.

K

155. Kahane JC. Connective tissue changes in the larynx and their effect on voice. *J. Voice* 1987; 1: 27-30.

156. Karnell MP, Scherer RS, Fischer LB. Comparison of acoustic voice perturbation measures among three independent voice laboratories. *J speech Hear Res*, 1991;34: 781-790.
157. Karnell M, Hall K, Landahl K. Comparison of fundamental frequency and perturbation measurements among three analysis systems. *J Voice* 1995;9: 383-393.
158. Kasuya H, Ogawa S, Mashima K, Ebihara S. Normalized noise energy as an acoustic measure to evaluate pathologic voice. *J Acoust Soc Am* 1986;80: 1329-1334.
159. Kay Elemetrics Corp. Multi Dimensional Voice Program (M.D.V.P.) Model 4305, Operations manual, Kay Elemetrics Corp. Pine Brook, 1993.
160. Kay Elemetrics. Disordered voice data base of the Massachusetts eye and ear infirmary voice and speech lab. Kay Elemetrics 2 Bridge Water Lane Lincoln Park, NJ07035-1488.
161. Kempster GB, Kistler DJ, Hillenbrand J. Multidimensional scaling analysis of dysphonia in two speaker groups. *J Speech Hear Res* 1991 ;34 : 534-543.
162. Kotby MN, Orabi AA. Voice range profile as a quantitative measure of vocal function in some pathological voices. In: Kotby MN (Ed.). *Congress Proceedings of XXIII IALP Congress, Cairo, August 1995, Cairo, IALP*, 1995.
163. Killian NJ. Perceptions of the voice-change process: Male adult versus adolescent musicians and non musicians. *J Research in Music Education*, 1997; 45: 521-535.
164. Kitzing P, Akerlund L. Long-time average spectrograms of dysphonic voices before and after therapy. *Folia Phoniatr* 1993;45: 53-61.
165. Klingholz F, Martin F. Die quantitative auswertung der stimmungsfeldmessung. *Sprache-Stimme-Gehör* 1983;7: 106-110.
166. Klingholz F, Martin F. Quantitative spectral evaluation of shimmer and jitter. *J Speech Hear Res* 1985;28: 169-174.
167. Kojima H, Gould W, Lambiase A, Isshiki N. Computer analysis of hoarseness. *Acta Otolaryngol* 1980; 89: 547-554.
168. Koufman J, Blalock D. Vocal fatigue and dysphonia in the professional voice user: Bogart-Bacall syndrome. *Laryngosc* 1988;98: 493-8.
169. Kreiman J, Gerratt BR, Precoda K. Listener experience and perception of voice quality. *J Speech Hear Res* 1990 ;33 : 103-15.

170. Kreiman J, Gerratt BR, Precoda K, Berke GS. Individual differences in voice quality perception. *J Speech Hear Res* 1992;35: 512-20.
171. Kreiman J, Gerratt BR, Kempster GB, Eрман A, Berke GS. Perceptual evaluation of voice quality: review, tutorial, and a framework for future research. *J Speech Hear Res* 1993;36: 21-40.
172. Kreiman J, Gerratt BR, Berke G. The multidimensional nature of pathologic vocal quality. *J Acoust Soc Am* 1994 ;96 : 1291-1302.

L

173. Ladefoged P. Some physiological parameters in speech. *Lang Speech* 1963;6: 109-119.
174. Lane H, Tranel B. The Lombard sign and the role of hearing in speech. *J Speech Hear Res* 1971;14: 677-709.
175. Large J, ed. *Vocal registers in singing*. The Hague: Mouton, 1973.
176. Lass N, Tecca J, Mancuso R, Black W. The effect of phonetic complexity on speaker race and sex identification. *J Phonetics* 1979;7:105-118.
177. Laver J. *The phonetic description of voice quality*. Cambridge University Press, London, 1980.
178. Laver J, Hiller S, MacKenzie J, Rooney E . An acoustic screening system for the detection of laryngeal pathology. *J Phonet* 1986 ;14 : 517-524.
179. Lazarus RS, *Stress coping and illness*. In: Friedman HS, ed. *Personality and disease*. New York: Wiley, 1990.
180. Leino T. Long-term average spectrum study on speaking voice quality in male actors. In: Friberg A, Iwarsson J, Jansson E, Sundberg J, eds. *SMAC 93, Proceedings of the Stockholm Music Acoustics Conference, July 28-August 1, 1993*. Stockholm: Royal Swedish Academy of Music, 1993: 206-210.
181. Leinonen L, Hiltunen T, Laakso ML, Rihkanen H, Poppius H. Categorization of voice disorders with six perceptual dimensions. *Folia Phoniatr Logop* 1997 ;49 : 9-20.
182. Lieberman P, Knudsen R, Mead J. Determination of the rate of change of fundamental frequency with respect to subglottal air pressure during sustained phonation. *J Acoust Soc Am* 1969;45: 1537-1543.

183. Lillge W: Die Reproduzierbarkeit von Stimmfeldern und deren Korrelation zur Tonhaltedauer, Inauguraldiss, Mainz, 1984, pp1-12.
184. Lin E, Jiang J, Noon SD, Hanson DG. Effects of head extension and tongue protrusion voice perturbation measures. *J Voice* 2000;14, 1: 8-16.
185. Lindblom B, Sundberg J. Acoustical consequences of lip, tongue, jaw, and larynx movement. *J Acoust Soc Am* 1971;50: 1166-1179.
186. Löfquist A, Kitzing F, Carlborg A Initial validation of an indirect measure of subglottal pressure during voicing. *J Acoust Soc Am* 1982, 72: 633-635.
187. Luchsinger R, Arnold G. *Voice Speech-language*. Belmont. CA: Wadsworth Publishing Co., 1965.

M

188. Mans EJ, Kuhn AG, Lamprecht-Dinnesen A, Psychosomatic findings in patients with vocal cord contact granulomas: initials results. *HNO* 1992; 40: 346-351.
189. Martin D, Fitch J, Wolfe V. Pathologic voice type and the acoustic prediction of severity. *J Speech Hear Res* 1995;38: 765-771.
190. Maxwell DE. A study of the effect of white noise masking on trained singers [Dissertation]. Norman, Oklahoma: University of Oklahoma, 1975.
191. Mayo R, Manning WH. Vocal tract characteristics of African-American and European-American adult males. *Tex J audiol Speech Pathol* 1994; 20: 33-36.
192. Mayo R, Grant WC. Fundamental frequency, perturbation, and vocal tract resonance characteristics of African-American and white American males. *Journal of the National Black association for Speech-language and Hearing* 1995;17: 32-38.
193. McAllister A, Sundberg J. Data on subglottal pressure and SPL at varied vocal loudness and pitch in 8- to 11-year-old children. *J Voice* 1998 ;12 : 166-174.
194. McGlone R, Hollien H. Vocal pitch characteristics of aged women. *J Speech Hear Res* 1963;6: 164-170.
195. McHenry M, Minton JT, Kuna ST, Vanoye CR, Robert-Seibert NS. Comparison of direct and indirect calculations of laryngeal airway resistance in various voicing conditions. *Europ J Disord Comm* 1995 ;30 : 435-449.

196. Mc Hugh-Munier C, Scherer KR, Lehmann W, Scherer U. Coping strategies, personality, and voice quality in patients with vocal fold nodules and polyps. *J Voice* 1997 ;11 : 452-461.
197. Meire Ph. Vie affective, psychodynamique et vieillissement. In Gommers A & van den Bosch de Aguilar, Ph, (eds). *Pour un vieillissement autonome*. Bruxelles: Mardaga Ed. 1992: 127.
198. Miller DG, Schutte HK. Feedback from spectrum analysis applied to the singing voice. *J Voice* 1990 ;4 : 329-334.
199. Miller DG, Schutte HK. Physical definition of the "flageolet register", 1993; 7, 3: 206-212.
200. Miller R. *The structure of singing*, Schirmer Books, New York, 1986.
201. Millet B, Dejonckere PH. What determines the differences in perceptual rating of dysphonia between experienced raters? *Folia Phoniatr Logop* 1998;50: 305-310.
202. Mishara BL, Riedel RG. Introduction au processus de vieillissement. In Mishara BL & Riedel RG (eds) *Le vieillissement*. Paris: Presses Universitaires de France, 1984: 11-28.
203. Morris RJ, Brown WS, Hicks DM, Howell E. Phonation profiles of male trained singers and nonsingers. *J Voice* 1995 ;9 : 142-148.
204. Morrison MD, Rammage LA, Gilles MB, Pullan CB, Nichol H, Muscular tension dysphonia. *The transactions of the American Laryngological association* 1983;104: 100-103.
205. Morsomme D, Remacle M, Millet B, High-resolution frequency analysis as applied to the singing voice. *Folia Phoniatica* 1993;45: 245-287.
206. Morsomme D, Boucquey D, Jamart J, Remacle M. Voice differences between the sexes in the elderly - comparison by Maximum Phonation Time, Phonation Quotient and Spectral Analysis. *Log. Phon. Vocol* 1997; 22: 9-14.
207. Morsomme D, Orban A, Jamart J, Remacle M. A study of vibrato: assessment by a panel of judges compared to spectral voice analysis. *Rev. Laryngol. Otol. Rhinol.* 1999; 120, 4:239-244.
208. Morsomme D, Jamart J, Wéry C, Giovanni A, Remacle M, Comparison between the GIRBAS Scale and EVA workstation for the assessment of dysphonia following unilateral vocal fold paralysis. *Folia Phoniatr. Logop.* (accepted for publication).

209. Murry T, Singh S, Sargent M : Multidimensional classification of abnormal voice qualities. J Acoust Soc Am 1977; 61: 1630-1635.
210. Murry T. Speaking fundamental frequency characteristics associated with voice pathologies. J Speech Hear Res 1978: 374-379.
211. Murry T, Doherty ET. Selected acoustic characteristics of pathologic and normal speakers. J Speech Hear Res 1980;23: 361-369.

N

212. Nawka T, Anders LC, Cebulla M, Zurakowski D. The speaker's formant in male voices. J Voice 1997 ;4 : 422-428.
213. Nemec J. The motivation background of hyperkinetic dysphonia in children: a contribution to psychologic research in phoniatriy. Logos 1961; 4: 28-31.
214. Netsell R. Subglottal and intraoral air pressures during the intervocalic contrast of /t/ and /d/. Phonetica 1969;20: 68-73.
215. Netsell R, Lotz WK, Peters JE, Schutte L. Developmental patterns of laryngeal and respiratory function for speech production. J Voice, 1994, 8: 123-131.
216. Neuschaefer-Rube Ch, Sram F, Klajman S. Three-dimensional phonetographic assessment of voice performance in professional and non-professional speakers. Folia Phoniatri logop, 1997; 49: 96-104.
217. Novak a, Vokral. Acoustic parameters for the evaluation of voice of future voice professionals, folia Phoniatri logop, 1995; 47: 279-285.

O

218. Ohala J. Respiratory activity in speech: In: Hardcastle W, Marchal A, eds speech Production and Speech Modelling. NATO ASIseries Behavioural and social sciences, 1990, 55: 23-53.
219. Ohman S, Lindqvist J. Analysis-by-synthesis of prosodic pitch contours. Z Phonet 1968;21: 164-170.
220. Olson DE, Goding GS, Michael DD. Acoustic and perceptual evaluation of laryngeal reinnervation by ansa cervicalis transfer. Laryngosc 1998 ;108 : 1767-1772.

221. Omori K, Kacker A, Carroll LM, Riley WD, Blaugrung SM. Singing power ratio : quantitative evaluation of singing voice quality. J Voice 1996 ;10 : 228-235.
222. Orlikoff R. The atherosclerotic voice. Ear Nose Throat J 1990;69: 833-837.
223. Orlikoff RF, Kahane JC. Influence of mean sound pressure level on jitter and shimmer measures. J Voice 1991 ;5 :113-119.
224. Osgood CE, Suci GJ, Tannenbaum PH. The measurement of meaning. University of Illinois Press, Urbana, 1957.

P

225. Pabon JPH, Plomp R. Automatic phonetogram recording supplemented with acoustical voice quality parameters. J Speech Hear Res. 1988;31: 710-722.
226. Pabon JPH. Objective voice-quality parameters in the computer phonetogram. J Voice 1991;5: 203-216.
227. Papcun G, Kreiman J, Davis A. Long-term memory for unfamiliar voices. J Acoust Soc Am 1989; 85 : 913-925.
228. Pedersen M. A longitudinal pilot study on phonetograms/voice profiles in pre-pubertal choir boys. Clin Otolaryngol 1993;18: 488-491.
229. Peppard R, Bless D, Milenkovic P. Comparison of young adult singers and nonsingers with vocal nodules. J Voice 1988;2: 250-260.
230. Perry CK, Ingrisano DR, Scott SR. Accuracy of jitter estimates using different filter setting on visi-pitch : a preliminary report. J Voice 1996 ;10 : 337-341.
231. Piccirillo JF, Fuller D, Painter C, Fredrickson JM. Multivariate analysis of objective vocal function. Ann Otol Rhinol Laryngol 1998 ;107 : 107-112.
232. Piccirillo JF, Painter C, Haiduk A, Fuller D, Fredrickson JM. Assessment of two objective voice function indices. Ann Otol Rhinol Laryngol 1998 ;107 : 396-400.
233. Pierce J. Le son musical : musique, acoustique et informatique. Pour la science. Paris, 1984.
234. Plant RL, Hillel AD, Waugh PF. Analysis of voice changes after thyroplasty using linear predictive coding. Laryngosc 1997 ;107 : 703-709.
235. Plant R, Hillel A. Direct measurement of subglottic pressure and laryngeal resistance in normal subjects and spasmodic dysphonia, J Voice, 1998; 12: 300-314.
236. Poburka BJ. A new stroboscopy rating form. J Voice, 1999;13: 403-413.

237. Prater R, Swift R. Manual of voice therapy. Austin, TX : Pro-ed, 1984.
238. Ptacek Ph, Sander EK. Maximum duration of phonation. J Speech Hear Res, 1963; 28: 171-182.
239. Ptacek Ph, Sander EK, Maloney WH, Jackson CCR. Phonatory and related changes with advanced age. J Speech Hear Res, 1966;9: 353-360.

Q

240. Quanjer PH. Standardized lung function testing. Report working party standardization of lung function tests, (ed) European Community for Coal and Steel. Bull. Europ. Physiopath. Respir. 1983;5: 1-95.

R

241. Rabinov CR, Kreiman J, Gerrat BR, Bielałowicz S. Comparing reliability of perceptual ratings of roughness and acoustic measures of jitter. J Speech Hear Res 1995 ;38 : 26-32.
242. Raes JPF, Clement PAR. Aerodynamic measurements of voice production. Acta O.R.L. Belg 1996;50: 293-298.
243. Ramig L, Ringel R. Effects of physiological aging on selected acoustic characteristics of voice. J Speech Hear Res 1983;26: 22-30.
244. Ramos C, Bless DM, Harmon R, Ford C. The mucosal wave as a prognostic sign in vocal paralysis. Paper presented at the meeting of the American Speech-language-Hearing Association; November 1993; Anaheim, Calif.
245. Rantala L, Määttä T, Vilkman E. Measuring voice under teacher's working circumstances: Fundamental perturbation features in maximally sustained phonation. Folia Phoniatr Logop 1997; 49: 281-291.
246. Reid C, The free voice (2nd ed). The Joseph Patelson Music House, New York, 1972.
247. Remacle M, Van Heule P, Ryckaert MA, Duvivier D, Bertrand B. Application clinique de l'analyse spectrale moyenne de la voix. Acta O.R.L. Belg 1988;42: 84-91.
248. Remacle M, Millet B, Van den Eeckhaut J. Application de l'analyseur de fréquence à haute résolution. Proposition de standardisation. Comptes-rendus

- des scéances, 86^{ème} congrès français d'O.R.L. et de pathologie cervico-faciale. Arnette ed., 1989: 159-64.
249. Remacle M, Millet B, Van Heule P, Duvivier D. Clinical application of the high-resolution frequency analyzer. *Folia Phoniatr* 1989;41: 259-67.
 250. Remacle M, Millet B, Duvivier D, Van Heule P. Application de l'analyse de fréquence à haute résolution en pathologie de la corde vocale. *Rev Fr Soc Franç O.R.L.* 1990;6: 15-24.
 251. Remacle M, Millet B. Etude objective de la qualité de la voix après laryngectomie partielle, Bruxelles, Ed. Acta Otot-rhino-laryngol Belg., 1991.
 252. Remacle M, Trigaux I. Characteristics of nodules through the high resolution frequency analyser, *Folia Phoniatr*, Basel, 1991; 43: 53-59.
 253. Remacle M, Dujardin JM, Lawson G. Treatment of vocal fold immobility by glutaraldehyde-cross-linked collagen injection: long-term results. *Ann Otol Rhinol Laryngol* 1995; 104: 437-441.
 254. Remacle M. The contribution of videostroboscopy in daily ENT practice. *Acta O.R.L. Belg* 1996;50: 265-281.
 255. Remacle M, Lawson G. Exploration du larynx. *Encyclop Méd Chir (Elsevier, Paris), Oto-rhino-laryngologie*, 20-635-A-10, 1997, p10.
 256. Revis J, Giovanni A, Wuyts F, Triglia JM. Comparison of different voice samples for perceptual analysis. *Folia Phoniatr Logop* 1999;51: 108-116.
 257. Roberts M, The pronunciation of vowels in negro speech. State University, Columbus, OH. Dissertation. 1966.
 258. Rohen JW, Yokochi C, Anatomie humaine. Atlas photographique d'anatomie systématique et topographique, Edition Vigot, (2^{ème}), Paris, 1991, pp 472.
 259. Rothenberg M, Mahshie J. Induced transglottal pressure variation during voicing. *Phonetics* 1986;14: 365-371.
 260. Rothman HB, Arroyo A. Acoustic variability in vibrato and its perceptual significance. *J Voice*, 1987; 4 : 123-141.
 261. Rothman HB, Rullman JF, Arroyo AA. Inter-and intra-subject changes in vibrato: Perceptual and acoustic aspects. *J Voice* 1990;4: 309-316.
 262. Rubin HJ, LeCover M, Vennard W. Vocal intensity, subglottic pressure and airflow relationships in singers. *Folia Phoniatr* 1967 ;19 : 393-413.

263. Sakata T, Kubota N, Yonekawa H, Imaizumi S, Niimi S. GRBAS evaluation of running speech and sustained phonation. In: Kotby N, ed. Proceedings of the world congress of the XXII IALP, Cairo, 6-10, August 1995. IALP, Cairo, 1995: 33-36.
264. Sander EK, Ripich DE. Vocal fatigue. *Ann Otol Rhinol Laryngol* 1983, 92: 141-145.
265. Sapienza CM. Aerodynamic and acoustic characteristics of the adult African American voice. *J Voice*. 1997; 11,4: 410-416.
266. Sapir S, Larson KK. Supralaryngeal muscle activity during sustained vibrato in four sopranos: Surface EMG Findings. *J Voice* 1993;7: 213-218.
267. Sataloff RT. Rational thought : the impact of voice science upon voice care. *J Voice* 1995 ;9 : 215-234.
268. Sataloff RT, Rosen DC, Hawkshaw M; Spiegel JR. The aging adult voice. *J Voice* 1997;11: 156-160.
269. Sataloff TR, MD, DMA. Professional Voice. The science and art of clinical care (2nd ed) Singular Publishing Group, Inc San Diego California 1997:
270. Sawashima M, Hirano M. Clinical evaluation of voice disorders. *Annu Bull Res Inst Logop Phoniater (Tokyo)* 1981;15: 165-171.
271. Sawashima M, Niimi S, Horigushi s, Yamaguchi H. On the clinical evaluation of aerodynamic data using the airway interruption method. *Ann Bull RILP*, 1987; 21: 101-110.
272. Scherer KR. Nonlinguistic vocal indicators of emotion and Psychopathology. In: Valdman A, ed. *Emotions in personality and Psychopathology*. New York, NY: Academic Press, 1979: 493-529.
273. Scherer KR. Speech and emotional states. In: Darby JK, ed. *Speech evaluation in psychiatry*. New York: Grune & Stratton, 1981: 189-220.
274. Scherer KR. Vocal assessment of affective disorders. In Maser JD, ed. *Depression and expressive behavior*. Hillsdale, NJ: Erlbaum, 1987: 57-83.
275. Scherer MC, Titze IR, Raphael BN, Wood RP, Ramig LA, Blager RF. Vocal fatigue in a trained and an untrained voice user. In : Baer T, Sasaki C, Harris K, eds. *Laryngeal function in phonation and respiration*. San Diego : Singular, 1991.

276. Schneider CM, Dennehy CA, Saxon KG. Exercise physiology principles applied to vocal performance : the improvement of postural alignment. *J Voice* 1997 ;11 : 332-337.
277. Schoentgen J. Jitter in sustained vowels and isolated sentences produced by dysphonic speakers. *Speech Comm* 1989 ;8 : 61-79.
278. Schultz-Carlon HJ, Bestimmung und beurteilung der individuellen mittleren spechstimmlage. *Folia Phoniatr.*, 1975; 27: 375-386.
279. Schutte H. The efficiency of voice production. San Diego, CA: Singular Publishing Group, 1980.
280. Schutte HK, Seidner W. Recommendation by the Union of European Phoniaticians (UEP): Standardizing voice area measurement/phonetography. *Folia Phoniatr.*, 1983;35: 286-288.
281. Scotto Di Carlo N. Internal voice sensitivities in opera singers. *Folia Phoniatr Logop* 1994 ;46 : 79-85.
282. Seashore CE. *Psychology of music*, New York: Dover Publications, Inc. 1937.
283. Siegwart H, Scherer KR. Acoustic concomitants of emotional expression in operatic singing : the case of Lucia in *Ardi Gli Incensi*. *J Voice* 1995 ;9 : 249-260.
284. Sindo ML, Hanson DG. Geriatric voice and laryngeal dysfunction. *Otolaryngol Clin North Am* 1990;23: 1035-1044.
285. Slis IH. Articulatory measurements on voiced, voiceless, and nasal consonants. *Phonetica* 1970;21: 193-210.
286. Smitheran JR, Hixon TJ. A clinical method for estimating laryngeal airway resistance during vowel production, *J. Speech Hear. Dis*, 1981; 46: 138-146.
287. Snidecor JC. The pitch and duration characteristics of superior female speakers during oral reading. *Journal of Speech and Hearing Disorders*, 1951;16: 44-52.
288. Södersen M, Hertegard S, Hammarberg B. Glottal closure transglottal airflow and voice quality in healthy middle-aged women. *J Voice* 1995; 9, 2: 182-197.
289. Sonninen A, Hurme P. Vocal fold strain and vocal pitch in singing : radiographic observations of singers and non-singers. *J Voice* 1998 ;12 : 274-286.
290. Sonninen A, Hurme P, Laukkanen A-M. The external frame function in the control of pitch, register and singing mode: Radiographic observations of a male singer. *J Voice*, 1999;13, 3: 319-340.

291. Sorensen D, Horii Y. Frequency and amplitude perturbation in the voices of female speakers. *J Commun Dis* 1983;16: 57-61.
292. Sorenson DN. A FO. Investigation of children ages 6-10 years old. *J Commun Dis* 1989;22: 115-23.
293. Soremed. Instrumentation médicale scientifique, E.V.A., Evaluation vocale assistée, manuel condensé d'utilisation des logiciels. Laboratoire parole et Langage, Marseille, 1993.
294. Spencer LE. Speech characteristics of male-to-female transsexuals: a perceptual and acoustic study. *Folia Phoniatr* 1988;40: 31-42.
295. Spirodouso WW. Physical fitness in relation to motor aging. IN: Mortimer JA, Pirozzolo FJ, Maletta GJ, eds. *The aging motor system*. New York: Praeger Publishers, 1982.
296. SQ-lab. E.V.A. 2 Mode d'emploi.
297. Stathopoulos ET, Weismer G. Oral airflow and air pressure during speech production: a comparative study of children, youths, and adults. *Folia Phoniatr* 1985;37: 152-159.
298. Stathopoulos ET, Sapienza C. Respiratory and laryngeal measures of children during vocal intensity variation. *J Acoust Soc Am* 1993;94(5): 2531-2543.
299. Stemple JC, Stanley J, Lee L. Objective measures of voice production in normal subjects following prolonged voice use. *J Voice* 1995 ;9 : 127-133.
300. Subtelney JD, Worth JH, Sakuda M. Intraoral air pressure and rate of flow during speech. *J Speech Hear Res* 1966;9: 498-518.
301. Sulter AM, Schutte HK, Miller DG. Differences in phonetogram features between male and female subjects with and without vocal training. *J Voice* 1995; : 363-377.
302. Sulter AM, Schutte HK, Miller DG. Standardized laryngeal videostroboscopic rating : differences between untrained and trained male and female subjects, and effects of varying sound intensity, fundamental frequency, and age. *J Voice* 1996 ;10 : 175-189.
303. Sundberg J. Formant structure and articulation of spoken and sung vowels. *Folia Phoniatr*, 1970, 22: 28-48.
304. Sundberg J. Formant technique in a professional soprano. *Acoustica* 1975 ;32 : 89-96.
305. Sundberg J. The acoustics of the singing voice. *Sc Am* 1977 ;236(3) : 82-91.

306. Sundberg J. The science of the singing voice. Dekalb : Northern Illinois University Press, 1987.
307. Sundberg J. Chest wall vibrations in singers. J Res Singing 1990;13: 25-54.
308. Sundberg J: Vocal Tract Resonance: the Science and Clinical Care. New York, Raven Press, 1991, chapter 4: 49-68.
309. Sundberg J, Titze IR, Scherer RC. Phonatory control in male singing : a study of the effects of subglottal pressure, fundamental frequency, and mode of phonation on the voice source. J Voice 1993 ;7 : 15-29.
310. Sundberg J, Elliot N, Gramming P, Nord L. Short-term variation of subglottal pressure for expressive purposes in singing and stage speech : a preliminary investigation. J Voice 1993 ;7 : 227-237.
311. Sundberg J. Perceptual aspects of singing. J Voice 1994 ;8 : 106-122.
312. Sundberg J, Skoog J. Dependence of jaw opening on pitch and vowel in singers. J Voice 1996 ;11: 301-306.
313. Sundberg J, Skoog J. Dependence of jaw opening on pitch and vowel in singers, 1997; 3: 301-306.
314. Swine Ch. Le vieillissement normal et pathologique. Notes de cours. Ciaco Bruxelles: Ciaco, 1992: 6-23.
315. Swine Ch. Caractéristiques de la personne âgée malades et problèmes fréquents en pathologie gériatrique. Louvain Médical 1992; 111: 209-214.

T

316. Tanaka S, Hirano M, Chijiwa K. Some aspects of vocal fold bowing. Ann Otol Rhinol Laryngol 1994;103: 357-62.
317. Taylor JT, Johnson TS. Voice disorders: The measurement of clinical progress. In Costello JM (ed.) Speech disorders in adults. Recent advances. San Diego: College-Hill Press, 1985: 127-152.
318. Teitler N. Examiner bias : influence of patient history on perceptual ratings of videostroboscopy. J Voice 1995 ;9 : 95-105.
319. Timiras PS. Biological perspectives on aging. Am Sci 1978; 68: 605-612.
320. Titze IR, Horii Y, Scherer R. Some technical considerations in voice perturbation measurements. J Speech Hear Res 1987 ;30 : 252-260.
321. Titze IR. Framework for the study of vocal register. J Voice 1988;2: 183-94.

322. Titze IR. On the relation between subglottal pressure and fundamental frequency in phonation. *J Acoust Soc Am* 1989;85: 901-906.
323. Titze IR, Luschei ES, Hirano M. Role of thyroarytenoid muscles in regulation of fundamental frequency. *J Voice* 1989;3: 213-224.
324. Titze IR. Phonation threshold pressure: a missing link in glottal aerodynamics. *J Acoust Soc Am* 1992;91: 2926-2935.
325. Titze IR, Sundberg J. Vocal intensity in speakers and singers. *J Acoust Soc Am* 1992;91: 2936-2946.
326. Titze IR, Liang H. Comparison of Fo extraction methods for high-precision voice perturbation measurements. *J Speech Hear Res* 1993 ;36 : 1177-1190.
327. Titze IR, Liang H. Principles of voice production. Prentice Hall. Englewood Cliffs 1994.
328. Titze IR. Toward standards in acoustic analysis of voice. *J Voice* 1994 ;8 : 1-7.
329. Titze IR. Workshop on acoustic voice analysis. Summary statement. Iowa-City: National center for voice and speech, 1995.
330. Tonkinson S. The Lombard effect in choral singing. *J Voice*, 1994; 8, 1, 24-29.
331. Trullinger RW, Emanuel FW. Airflow characteristic of stop-plosive consonants productions of normal-speaking children. *J Speech Hear Res* 1983;26: 202-208.

V

332. Van de Heyning Ph, Remacle M, Van Cauwenberge P. Functional assessment, of voice disorders. *Acta oto-rhino-laryngol belg.*, 1996; 50.
333. Van den Berg J, Vennard W. Toward an objective vocabulary. *Nats Bull* 1959; 15: 10-15.
334. van den Berg JW. Subglottal pressure and vibration of the vocal folds. *Folia Phoniater* 1957;9: 65-71.
335. Van Lierde K, Moermen M, Vermeersch H, Van Cauwenberge P. An introduction to computerised speech lab. *Acta O.R.L. Belg* 1996;50: 109-114.
336. Vennard W. Singing: the mechanism and the technique. New York: Carl Fischer, Inc., 1967.
337. Vilkman E, Lauri E-R, Paavo Alku, Sla E, Sihvo M. Effects of prolonged oral reading on Fo, SPL, subglottal pressure and amplitude characteristics of glottal flow wave forms, 1999; 13, 2, 305-315.

338. Vuorenkoski V, Lenko HL, Tjernlund O, Vuorenkoski L, Perheentupa J. Fundamental voice frequency during normal and abnormal growth, and after androgen treatment. *Arch Dis Child* 1978;53: 201-209.

W

339. Waar CH, Damsté PH. Het fonetogram. *Log Phoniater*. 1968;40: 198-201.
340. Walton J, Orlikoff R, Speaker race identification from acoustic cues in the vocal signal. *J Speech Hear Res* 1994;37: 738-745.
341. Wapnick J, Darrow AA, Kovacs J, Dalrymple L. Effects of physical attractiveness on evaluation of vocal performance. *JRME* 1997 ;45 : 470-479.
342. Wapnick J, Ekholm E. Expert consensus in solo voice performance evaluation. *J Voice* 1997 ;11 : 429-436.
343. Ward PH, Colton R, McConnell F, Malmgren L, Kashima H, Woodson G. Aging of the voice and swallowing. *Otolaryngol Head Neck Surg* 1989;100: 283-286.
344. Wheat MC, Hudson AI. Spontaneous speaking FO of 6-year-old black children. *J Speech Hear Res* 1988;31: 723-725.
345. Wilcox KA, Horii Y. Age and changes in vocal jitter. *J Gerontol* 1980;35: 194-198.
346. Wilson FB, Lamb MM. Comparison of personality characteristics of children with and without vocal nodules on Rorschach protocol interpretation. *Acta Symbolica* 1973;5: 43-55.
347. Wilson DK. Voice problems of children (2nd ed). Williams & Wilkins, Baltimore, 1979: 63-95.
348. Wilson DK. Voice problems of children (3rd ed). Williams & Wilkins, Baltimore, 1987.
349. Wirz S, Mackenzie Beck J. Assessment of voice quality: the vocal profiles analysis scheme: In: *Perceptual approaches to communicate disorders*. S Wirz (Ed) Whum Publisher, Inc. 1937.
350. Wolfe V, Steinfatt T. Prediction of vocal severity within and across voices type. *J Speech Hear Res* 1987;30: 230-240.
351. Wolfe VI, Ratusnik DL, Smith FH. Intonation and fundamental frequency in male-to-female transexuals. *J Speech Hear Res* 1990 ;55 : 43-50.

352. Wolfe V, Fitch J, Cornell R. Acoustic prediction of severity in commonly occurring voice problems. *J Speech Hear Res* 1995 ;38 : 273-9.
353. Wolfe V, Fitch J, Martin D. Acoustic measures of dysphonic severity across and within voice types. *Folia Phoniatr Logop* 1997 ;49 : 292-299.
354. Wolfe V, Martin D. Acoustic correlates of dysphonia : Type and severity. *J Comm Disord* 1997 ;30 : 403-416.
355. Woo P, Casper J, Colton R, Brewer D. Dysphonia in the aging: physiology versus disease. *Laryngosc* 1992;102: 139-144.
356. Wuyts FL, De Bodt MS, Molenberghs G, Remacle M, Heylen L, Millet B, Van Lierde C , Raes J, Van de Heyning PH. The dysphonia severity index : an objective measure of hoarseness based on a multi-parameter approach. *J Speech Hear Res* 2000;43: 1-13.

Y

357. Yanagihara N. Significance of harmonic changes and noise components in hoarseness. *J Speech Hear Res* 1967;10: 531-541.
358. Yoshida H. A video-tap recording system for laryngostroboscopy. *J Jap Bronchoesoph Soc* 1979;30: 6-12.
359. Yumoto E, Sasaki Y, Okamura H. Harmonics-to-noise ratio and psychophysical measurement of the degree of hoarseness. *J Speech Res* 1984;27: 2-6.

Z

360. Zajac DJ, Mayo R. Aerodynamic and temporal aspects of velopharyngeal function in normal speakers. *J Speech Res* 1996;39: 1199-1207.
361. Zajac DJ. Effects of a pressure target on laryngeal airway resistance in children. *Comm. Disord*, 1998; 31: 201-213.
362. Zeitels SM. Reinke's Edema: Phonatory mechanisms and management strategies. *Ann Otol-rhinol-laryngol.*, 1997; 107: 533-542.
363. Zemlin WR. *Speech and hearing science, anatomy and physiology*. 2nd ed. Englewood Cliffs, NJ: Prentice-Hall, 1981.

Tableaux

CHAPITRE 1.

Tableau 1 : Valeurs de la fréquence fondamentale d'enfants âgés de 3 à 10 ans^[8, 46, 80, 99, 116, 143, 292, 344].

Tableau 2 : Valeurs de la fréquence fondamentale de sujets âgés de 20 à plus de 65 ans^[29, 46, 203].

Tableau 3 : Moyennes fréquentielles et déviations standards en fonction de trois types d'intensité rapportées par Sulter et al (1996)^[302].

Tableau 4 : Valeurs de la fréquence fondamentale de chanteurs rapportées par Nadolesny (1923)^[in 69].

Tableau 5 : Valeurs de la fréquence fondamentale pour des chanteurs rapportées par Brown (1991)^[32].

Tableau 6 : Valeurs de la fréquence fondamentale pour des chanteurs et non-chanteurs rapportées par Morris et al (1995)^[203].

Tableau 7 : Valeurs de la fréquence fondamentale pour des chanteurs rapportées par Drew et al (1995)^[69].

Tableau 8 : Valeurs de la fréquence fondamentale pour des chanteurs rapportées par Novak et al (1995)^[217].

Tableau 9 : Valeurs de la fréquence fondamentale en fonction de la pathologie rapportées par Peppard (1988)^[229] et Drew et al (1995)^[69].

Tableau 10 : Valeurs de la fréquence fondamentale parlée^[6, 90, 294, 351].

Tableau 11 : Valeurs de la fréquence fondamentale en fonction de la posture de la tête et des mouvements linguaux rapportées par Lin et al (2000)^[229].

Tableau 12 : Valeurs de l'intensité selon différents auteurs^[33, 40, 89, 203].

Tableau 13 : Valeurs de jitter rapportées par Giovanni et al (1996)^[95].

Tableau 14 : Valeurs de jitter en fonction des méthodes de calcul d'extraction de la fréquence fondamentale rapportées par Titze (1993)^[326].

Tableau 15 : Valeurs de jitter selon différents auteurs^[15, 29, 80, 131, 138, 140, 153, 166, 211, 243, 291].

Tableau 16 : Corrélations entre les paramètres acoustiques et les paramètres perceptuels selon différents auteurs^[50, 55, 61, 77, 167, 189, 201, 350, 353, 354, 352, 359].

CHAPITRE 3.

Tableau 17 : Profil vocal de Laver, "Vocal Profile Analysis" (VPAS) De Bodt et al (1996)^[51].

Tableau 18 : Nombre de cycles par seconde du vibrato calculé par différents auteurs [38, 63, 107, 200, 260, 261, 326].

Tableau 19 : Critères pour une production vocale souhaitée et une production vocale défectueuse Van den Berg et al (1959)^[332].

Tableau 20 : Degrés de compétence, niveaux d'étude et tessiture des 9 sujets chanteurs [étude 1].

Tableau 21 : Paramètres choisis en fonction des différents tris [étude 1].

Tableau 22 : Pourcentage de concordance intra-juges pour le type d'oscillation [étude 1].

Tableau 23 : Pourcentage de concordance inter-juges pour le type d'oscillation tous juges confondus [étude 1].

Tableau 24 : Concordances intra-juges pour les 6 paramètres (non-gênant/gênant, harmonieux/dysharmonieux, agréable/désagréable, régulier/irrégulier, stable/instable, timbré/détimbré) [étude 1].

Tableau 25 : Concordances inter-juges pour les 6 paramètres (non-gênant/gênant, harmonieux/dysharmonieux, agréable/désagréable, régulier/irrégulier, stable/instable, timbré/détimbré) [étude 1].

Tableau 26 : Variabilité inter-juges pour les 6 paramètres dans les 3 tâches ([a] vibré, [a] non-vibré et morceau imposé) [étude 1].

Tableau 27 A : Concordances inter-juges pour les 9 chanteurs dans les 3 tâches, tous paramètres confondus (Déviations standards) [étude 1].

Tableau 27 B : Valeurs de p pour les déviations standards significatives entre les 6 paramètres [étude 1].

Tableau 28 : Types d'oscillation test/re-test. Comparaisons par rang toutes tâches et tous juges confondus [étude 1].

Tableau 29 : Résultats pour la tâche [a] vibré pour l'oscillation tous juges confondus [étude 1].

Tableau 30 : Résultats pour la tâche [a] non-vibré pour l'oscillation tous juges confondus [étude 1].

Tableau 31 : Résultats pour la tâche du morceau imposé pour l'oscillation tous juges confondus [étude 1].

Tableau 32 : Analyse en composante principale, paramètres x tâche ([a] vibré) [étude 1].

Tableau 33 : Analyse en composante principale : paramètres x tâche (morceau imposé) [étude 1].

Tableau 34 : Analyse en composante principale : paramètres x tâche ([a] non-vibré) [étude 1].

CHAPITRE 4.

Tableau 35 : Valeurs de temps maximum phonatoire rapportées par Raes et al^[242].

Tableau 36 : Valeurs de capacité vitale rapportées par Raes et al^[242].

Tableau 37 : Valeurs de quotient phonatoire rapportées par Raes et al^[242].

Tableau 38 : Niveaux de pression sonore en fonction de la hauteur tonale pour les filles (n=67) et les garçons (n=45) âgés entre 7 et 10 ans Böhme et al (1995)^[25].

Tableau 39 A : Niveaux de pression sonore et temps maximum phonatoire d'orateurs professionnels et non-professionnels rapportés par Neuschaefer et al (1997)^[216].

Tableau 39 B : Niveaux de pression sonore et temps maximum phonatoire d'oratrices professionnelles et non-professionnelles rapportés par Neuschaefer et al (1997)^[216].

Tableau 40 : Valeurs de pression sous-glottique en fonction de 3 méthodes. Auteurs cités par Giovanni (1998)^[98].

Tableau 41 : Valeurs de pression intra-orale et transglottique (cmH₂O) rapportées par McHenry (1995)^[195].

Tableau 42 : Valeurs moyennes de la pression sous-glottique estimée (Hpa) mesurées dans différentes conditions d'intensité (écart-types entre parenthèses) Giovanni (1998)^[98].

Tableau 43 : Valeurs moyennes de la pression sous-glottique (Hpa) mesurées en fonction de différentes hauteurs tonales (écart-types entre parenthèses) Giovanni (1998)^[98].

Tableau 44 : Valeurs de la pression sous-glottique (cmH₂O) en fonction de différentes tâches rapportées par Subtelney et al (1966)^[300], Bernthal et al (1978)^[20] et Stathopoulos et al (1993)^[298].

Tableau 45 : Valeurs de la pression sous-glottique pour les deux sexes en fonction des consonnes P, B, T, D rapportées par Subtelney (1966)^[300] et Bernthal et al (1978)^[20].

Tableau 46 : Valeurs de la pression sous-glottique estimée (cmH₂O) en fonction de l'intensité rapportées par Stathopoulos et al (1985)^[297] et Mac Allister (1998)^[193].

CHAPITRE 5.

Tableau 47 : Observation du formant du chanteur dans les 3 groupes pour chaque voyelle émise (/a/¹, /u/, /i/, /a/²) [étude 2].

Tableau 48 : Observation du vibrato dans les 3 groupes pour chaque voyelle émise (/a/¹, /u/, /i/, /a/²) [étude 2].

Tableau 49 : Observation du formant du chanteur et du vibrato en fonction de la tâche et du groupe (Test binomial) [étude 2].

Tableau 50 : Comparaison de l'acquisition du formant du chanteur et du vibrato en fonction des 3 groupes [étude 2].

Tableau 51 : Mesures extraites de l'étude d'Inglis et al (1985)^[144] et Taylor et al (1985)^[316] sur une population jeune [étude 3].

Tableau 52 : Valeurs moyennes de la capacité vitale, du temps maximum phonatoire, du quotient phonatoire et de la fréquence fondamentale pour les deux sexes [étude 3].

Tableau 53 : Comparaison des mesures quantitatives entre les deux populations [étude 3].

Tableau 54 : Valeurs de cadence et d'étendue du vibrato selon différents auteurs ^[38, in 63, 107, 200, 260, 261, 266, 327].

Tableau 55 : Concordance du juge par rapport à lui-même par tâche pour la question 1 concernant le type d'oscillation [étude 4].

Tableau 56 : Comparaison en pourcentage d'observations concordantes par tâche pour la question 1 concernant le type d'oscillation [étude 4].

- Tableau 57** : Concordances du juge 1 par rapport à lui-même par tâche pour la question 1 concernant le type d'oscillation [étude 4].
- Tableau 58** : Concordances des juges par rapport à eux-mêmes par juge pour la question 1 concernant le type d'oscillation [étude 4].
- Tableau 59** : Comparaison des pourcentages d'observations concordantes par juge pour la question 1 concernant le type d'oscillation [étude 4].
- Tableau 60** : Corrélations des juges par rapport à eux-mêmes par juge pour 4 variables et 4 tâches pour la question 2 concernant les paramètres [étude 4].
- Tableau 61** : Comparaison intra-juges 2 à 2 par tâche, en considérant 4 variables pour la question 2 concernant les paramètres [étude 4].
- Tableau 62** : Corrélations entre paramètres subjectifs et objectifs pour la question 1 concernant le type d'oscillation et le juge 1 [étude 4].
- Tableau 63** : Corrélations entre paramètres subjectifs et objectifs pour la question 2 concernant les paramètres et le juge 1 [étude 4].
- Tableau 64** : Statistiques descriptives réparties en fonction du groupe de sujets (non-chanteurs/chanteurs), de la tâche et du moment de l'émission (A, B, C, D) [étude 5].
- Tableau 65** : Valeurs pour les paramètres relatifs à la fréquence fondamentale [étude 5].
- Tableau 66** : Valeurs pour les paramètres relatifs à l'intensité [étude 5].
- Tableau 67** : Valeurs pour les paramètres relatifs au débit d'air buccal [étude 5].
- Tableau 68** : Valeurs pour les paramètres relatifs à la variabilité de la fréquence fondamentale : le jitter [étude 5].
- Tableau 69** : Valeurs pour les paramètres relatifs à la variabilité de l'intensité : le shimmer [étude 5].
- Tableau 70** : Valeurs pour les paramètres relatifs à la fuite glottique [étude 5].
- Tableau 71** : Concordances intra-juges aux 3 lectures [étude 6]
- Tableau 72** : Nombre de concordances entre les juges [étude 6]
- Tableau 73** : Analyse de variance des 6 paramètres de l'échelle G.I.R.B.A.S.
- Tableau 74** : Corrélations entre les paramètres de l'échelle G.I.R.B.A.S. (Spearman) [étude 6].
- Tableau 75** : Corrélations entre les paramètres subjectifs et objectifs (Spearman's Rho) [étude 6].

- Tableau 76 A** : Statistique descriptive des mesures acoustiques et aérodynamiques pour la tâche [a] simple [étude 7].
- Tableau 76 B** : Statistique descriptive du temps maximum phonatoire pour la tâche [a] tenu [étude 7].
- Tableau 76 C** : Statistique descriptive de la pression sous-glottique estimée pour la tâche [pa] [étude 7].
- Tableau 77** : Anova pour la fréquence fondamentale, l'intensité et le flux en fonction du sexe, de la tâche et de la répétition [étude 7].
- Tableau 78 A** : Statistique descriptive des mesures acoustiques et aérodynamiques pour la tâche [a] simple toutes pathologies confondues [étude 7].
- Tableau 78 B** : Statistique descriptive du temps maximum phonatoire pour la tâche [a] tenu, toutes pathologies confondues [étude 7].
- Tableau 78 C** : Statistique descriptive de la pression sous-glottique estimée pour la tâche [pa], toutes pathologies confondues [étude 7].
- Tableau 79 A** : Statistique descriptive des mesures acoustiques et aérodynamiques pour la tâche [a] simple, pour les sujets féminins porteurs de nodules [étude 7].
- Tableau 79 B** : Statistique descriptive du temps maximum phonatoire pour la tâche [a] tenu pour les sujets féminins porteurs de nodules [étude 7].
- Tableau 79 C** : Statistique descriptive de la pression sous-glottique estimée pour la tâche [pa] pour les sujets féminins porteurs de nodules [étude 7].
- Tableau 80 A** : Statistique descriptive des mesures acoustiques et aérodynamiques pour la tâche [a] simple, pour les sujets féminins porteurs d'un œdème de Reinke [étude 7].
- Tableau 80 B** : Statistique descriptive du temps maximum phonatoire pour la tâche [a] tenu pour les sujets féminins porteurs d'un œdème de Reinke [étude 7].
- Tableau 80 C** : Statistique descriptive de la pression sous-glottique estimée pour la tâche [pa] pour les sujets féminins porteurs d'un œdème de Reinke [étude 7].
- Tableau 81 A** : Statistique descriptive des mesures acoustiques et aérodynamiques pour la tâche [a] simple pour les sujets sains [étude 7].
- Tableau 81 B** : Statistique descriptive du temps maximum phonatoire pour la tâche [a] tenu pour les sujets sains [étude 7].

Tableau 81 C : Statistique descriptive de la pression sous-glottique estimée pour la tâche [pa] pour les sujets sains [étude 7].

Tableau 82 : Comparaison des mesures de fréquence fondamentale, d'intensité, de flux entre les 3 groupes de sujets féminins [étude 7].

Tableau 83 : Comparaison des mesures de fuite glottique, de S_r , de temps maximum phonatoire, d'APQ et de pression sous-glottique estimée entre les 3 groupes de sujets féminins.

Figures

CHAPITRE 1

Figure 1 : Boîtier ostéo-cartilagineux. Rohen et al (1991), p. 146^[258].

Figure 2 : Cartilages, membranes et ligaments du larynx. Stataloff et al (1997), p.112^[269].

Figure 3 : Muscles intrinsèques du larynx. Stataloff et al (1997), p.118^[269].

Figure 4 : Innervation du larynx. Rohen et al (1991), p. 150^[258].

Figure 5 : Muscles extrinsèques (vue latérale). Rohen et al, 1991, p. 144) [].

Figure 6 : Couches de la corde vocale. Stataloff et al (1997), p.115^[269].

Figure 7 : Vibration des cordes vocales en coupe frontale. Stataloff et al (1997), p.222^[269].

Figure 8 : Mouvement de bascule du cartilage thyroïde sur le cartilage cricoïde par l'action du muscle thyro-cricoïdien. Stataloff et al (1997), p.119^[269].

Figure 9 : Facteurs déterminant le spectre de la voyelle. Stataloff et al (1997)^[269].

Figure 10 : Formants des voyelles en voix chantée et parlée. Sundberg et al (1970)^[303].

Figure 11 : Formants du locuteurs et du chanteurs décrits par Sundberg et al (1970)^[303].

Figure 12 : Distribution en fréquences de la puissance pour un orchestre et pour le ténor : Jussi Björling. Pierce (1983), p. 126^[233].

Figure 13 a : Cordes vocales en coupe frontale et ondes glottiques illustrant les registres "pulse" et "loft" Stataloff et al (1997), p.155^[269].

Figure 13 b : Modèle vibratoire normale des cordes vocales Stataloff et al (1997), p.^[269].

CHAPITRE 2

Figure 14 : Cycles vibratoires. Images digitalisées sur écran. Remacle et al (1997), p. 5^[255].

Figure 15 : Principe de la stroboscopie. L'image fixe Remacle et al (1997), p. 5^[255].

Figure 16 : Principe de la stroboscopie. Le déphasage Remacle et al (1997), p. 5^[255].

Figure 17 : Principe de la stroboscopie. L'impression de ralenti Remacle et al (1997), p. 5^[255].

CHAPITRE 3

Figure 18 : Partition du morceau imposé pour la tessiture soprano, composé par G. Chapelier en 1995.

CHAPITRE 4

Figure 19 : Mesures des pentes du phonétogramme. Heylen et al (1996)^[118].

Figure 20 : Phonétogrammes en 3 dimensions d'orateurs et oratrices professionnels et non-professionnels. Neuschaefer et al (1997), p.99^[216].

Figure 21 : Analyseur dynamique de signal de Hewlett Packard modèle 3561A^[117].

Figure 22 : Schémas de l'analyse spectrale (A, B, C, D) illustrant 4 niveaux de l'échelle. Morsomme et al (1997),p.10-12^[207].

Figure 23 : Illustration du Computer Speech lab (CSL) de Kay elemetrics ^[159].

Figure 24 A : Dispositif E.V.A. (Evaluation vocale assistée) : boîtier électronique et micro-ordinateur. Photo 1998.

Figure 24 B : Dispositif E.V.A. : la pièce à main avec l'embouchure buccale en silicone souple. Photo 1998.

Figure 25 : Principe de l'estimation de la pression sous-glottique par l'intermédiaire de la pression intra-orale. SQ-lab ^[296].

Figure 26 : Schéma de la procédure de McHenry et al (1995)^[195].

CHAPITRE 5

Figure 27 A : Analyses spectrales d'un sujet du groupe 1(A) (étude 2). Morsomme et al (1993)^[282].

Figure 27 B-F: Formant du chanteur pour les 3 groupes (étude 2). Morsomme et al (1993)^[205].

Figure 28 : Distribution en pourcentage de la qualité vocale (obtenue à l'analyse spectrale et mesurée sur une échelle à 4 niveaux) des sujets âgés hommes et femmes (étude 3). Morsomme et al (1997)^[206].

Figure 29 : Mesures de la cadence moyenne du vibrato, de l'étendue moyenne et du délai de "mise en route" du vibrato. Les mesures sont obtenues d'un

spectrogramme à bande étroite à l'aide d'une procédure semi-automatique. Sataloff et al (1997), p. ^[269].

Figure 30 A : Première partie du questionnaire proposée au jury (étude 4). Morsomme et al (1999)^[287].

Figure 30 B : Deuxième partie du questionnaire proposée au jury (étude 4). Morsomme et al (1999)^[207].

Figure 31 : Comparaison entre les sujets dysphoniques et les sujets sains pour le temps maximum phonatoire (étude 7).

Figure 32 : Comparaison entre les sujets dysphoniques et les sujets sains pour la fuite glottique (étude 7).

Figure 33 : Comparaison entre les sujets dysphoniques et les sujets sains pour la pression sous-glottique estimée (étude 7).

Annexes

Annexe 1 : "The Rainbow passage"

Ce texte est lu à haute voix dans les évaluations vocales anglo-saxonnes, il contient tous les sons de la langue anglaise :

When the sunlight strikes raindrops in the air, they act like a prism and form a rainbow. The rainbow is a division of white light into many beautiful colors. These take the shape of a long round arch, with its path high above, and its two ends apparently beyond the horizon. There is, according to legend, a boiling pot of gold at one end. People look but no one ever finds it. When a man looks for something beyond his reach, his friends say he is looking for the pot of gold at the end of the rainbow.

Annexe 2 : Protocole du B.S.G.V.D.

VOICE DISORDERS - BELGIAN STUDY GROUP 1996

IDENTIFICATION

Centre ☐ UZA ☐ UZG ☐ UCL ☐ VUB ☐ TH Date ____/____/199...

Type of study ☐ normative ☐ Belgian Study Group ☐

Patient identification : name _____ Examiner identification (optional)
date of birth (YMD) _____ Dr. _____
Speech therapist _____

Type ☐ new patient ☐ known patient (in file)

Age (in years) at moment of registration : _____

Sex ☐ male ☐ female

Profession ☐ elite vocal performer (professional singers and actors...)
☐ professional voice user (lecturers, public speakers, telephone operators, clergy...)
☐ nonvocal professional (lawyers, business persons, sales persons, doctors, ...)
☐ nonvocal nonprofessional (factory worker, laborer, clerk ...)

CHARACTERISTICS OF THERAPY (for known patients only)

Control examination _____ (serial number)

Time between start of therapy and examination : ____ weeks ____ months ____ years

☐ no therapy ☐ surgery _____
☐ pharmacological _____ ☐ radiotherapy _____
☐ voice therapy _____ Therapy code (optional) : _____

1. Case history and medical diagnosis

Influencing factors ☐ vocal abusos ☐ smoking ☐ allergy ☐ infection U.R.T.
☐ C.O.L.D. ☐ nasal airway obstr. ☐ stress ☐ other _____

E.N.T. antecedents : _____

Clinical diagnosis and laryngoscopic findings
at moment of registration

Primary diagnosis : _____ severity 1 2 3

Secondary diagnosis : _____ severity 1 2 3

Severity rating 1 = minimal 2 = moderate 3 = severe

Diagnosis code

- | | |
|--|-----------------------------|
| 0. normal | 12. tumour |
| 1. excessive muscular tension/
no changes in laryngeal mucosa | 13. granuloma |
| 2. incompl. closure | 14. paralysis in abduction |
| no changes in laryngeal mucosa | 15. paralysis in adduction |
| 3. mucosal cyst | 16. spasmodic dysphonia |
| 4. vocal nodules | 17. other neurological dis. |
| 5. polyps | 18. laryngeal trauma |
| 6. epidermoid cyst | 19. endocrinological dis. |
| 7. sulcus glottidis | 20. ventricular phonation |
| 8. Reinke's oedema | 21. psychogenic aphonia |
| 9. vocal haemorrhage | 22. other : |
| 10. laryngitis-acute | 23. intermediate paralysis |
| 11. laryngitis-chronic | 24. scar |
| | 25. contact ulcer |
| | 26. vascular ectasia |

Diagnosis by means of

1.	Indirect laryngoscopy	0	1	2	3	4
2.a.	Rigid endoscopy	0	1	2	3	4
2.b.	Fiberoptic endoscopy	0	1	2	3	4
3.	Stroboscopy	0	1	2	3	4
4.	Videostroboscopy	0	1	2	3	4

Appreciation of examination
0 = not examined
1 = provides no new information
2 = usefull in diagnosing
3 = essential in diagnosing
4 = changed diagnosis

2. Stroboscopic examination

Fundamental frequency : ____ Hz

Symmetry	symmetrical		
	asymmetrical		
		in amplitude	
		in phase	
Regularity	regular		
	irregular		
	inconsistent		
Glottic closure	complete		
	incomplete	longitudinal	
		posterior	
		anterior	
		irregular	
		oval	
		hour-glass	
	inconsistent		
Amplitude		right	left
	increased		
	normal		
	reduced		
	none		
Mucosal wave	normal		
	reduced		
	none		
Lesion	nodule, oedematous		
	nod. fibro-oedematous		
	polyp		
	Reinke oedema I, II, III		
	sulcus and related lesions		
	mucosal cyst		
	keratosis		
	immobility		
	other :		

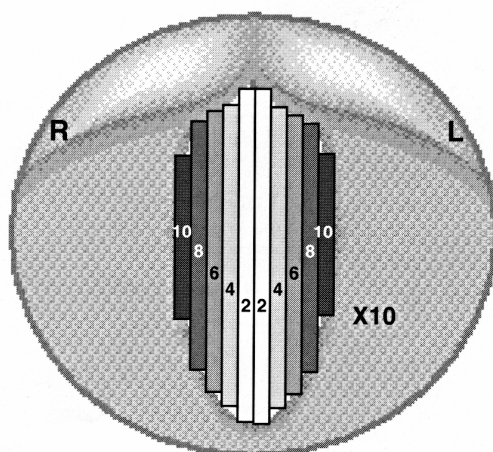
Annexe 3 : Protocole du S.E.R.F.

Stroboscopy Evaluation Rating Form (SERF)

Bruce J. Poburka, Ph.D

Rater: _____
Client: _____
Date: _____

Amplitude (Rate @ normal pitch & loudness)

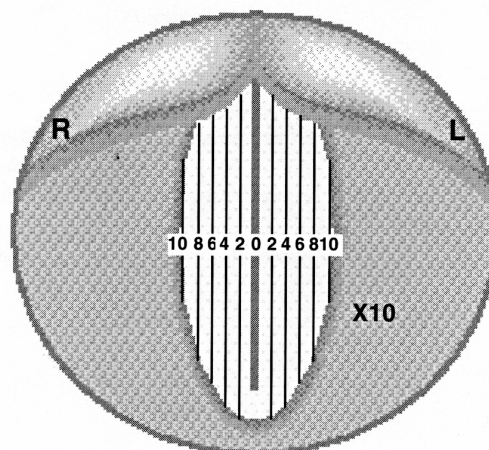


Right: _____%

Left: _____%

Fo: _____

Mucosal Wave (Rate @ normal pitch & loudness)

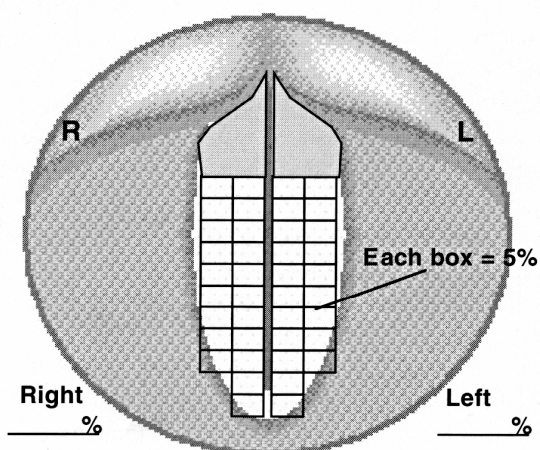


Right: _____%

Left: _____%

Fo: _____

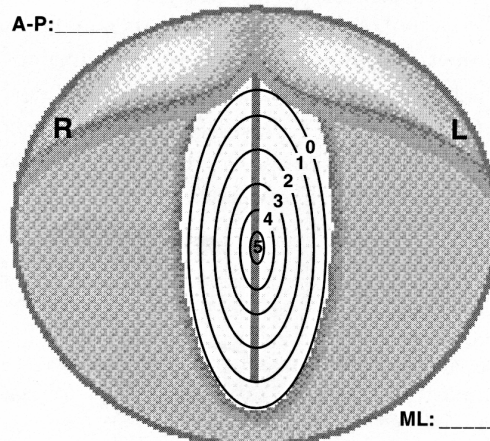
Non-vibrating Portion (shade in affected areas)



Right
_____%

Left
_____%

Supraglottic Activity (Ignore voice onsets)



A-P: _____

ML: _____

A

FIG. 1. A. The Stroboscopy Evaluation Rating Form (SERF). (continued)



Vocal Fold Edge Smoothness

Right Fold 0 1 2 3 4 5 smooth rough	circle on	Left Fold 0 1 2 3 4 5 smooth rough
--	-----------	---

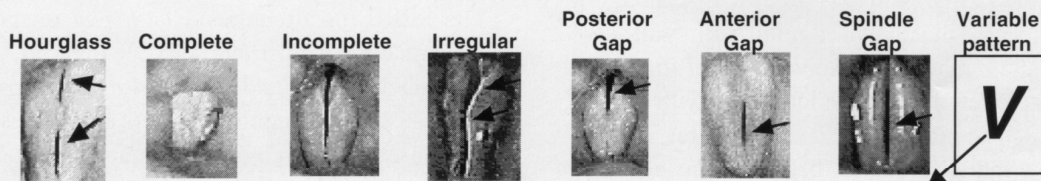
Vocal Fold Edge Straightness

Right Fold 0 1 2 3 4 5 straight irregular	circle on	Left Fold 0 1 2 3 4 5 straight irregular
--	-----------	---

Rate @normal pitch & loudness

Vertical Level	Phase Closure	Phase Symmetry	Regularity
circle one  on-plane	Rate @ point of contact % of time open closed Breathy +90% <10% circle one 66% 33% Normal 33% 66%	Rate @ point of contact % of time symmetrical Always asymmetrical circle one 0% 20% 40% 60% 80% 100%	% of time regular Always irregular circle one 0% 20% 40% 60% 80% 100%
 off-plane	<10% +90% Pressed / Fr Frame count: open phase: _____ Closed phase: _____	Always symmetrical	Always regular Method(s) used: stop phase _____ running phase _____

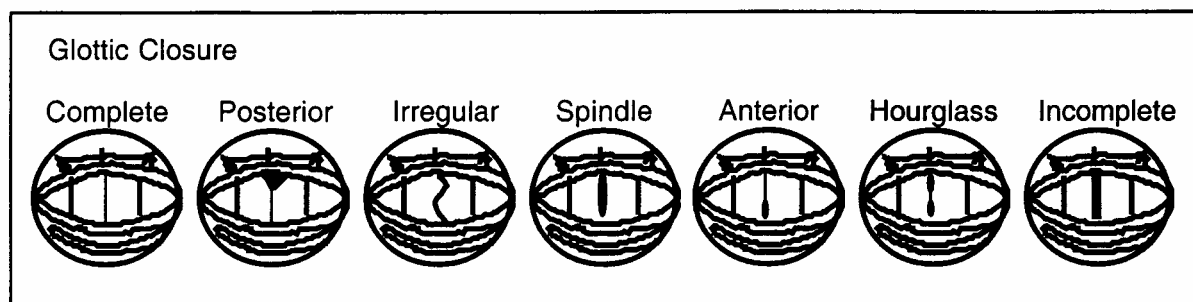
Glottal Closure



If closure pattern is variable, indicate the predominant closure pattern: _____

Summary/Additional Comments:

Annexe 4 : Protocole du S.A.V.



Scale for rating glottic closure on the SAV form. There is no provision for rating a variable glottal closure pattern.

Supraglottic Activity	①	②	③	④	⑤
	①	②	③	④	⑤
	①	②	③	④	⑤
	①	②	③	④	⑤
	①	②	③	④	⑤
	①	②	③	④	⑤
	①	②	③	④	⑤
	①	②	③	④	⑤
	①	②	③	④	⑤
	①	②	③	④	⑤
	①	②	③	④	⑤
	①	②	③	④	⑤
	①	②	③	④	⑤
	①	②	③	④	⑤
	①	②	③	④	⑤
	①	②	③	④	⑤
	①	②	③	④	⑤
	①	②	③	④	⑤
	①	②	③	④	⑤
	①	②	③	④	⑤
	①	②	③	④	⑤
	①	②	③	④	⑤
	①	②	③	④	⑤
	①	②	③	④	⑤
	①	②	③	④	⑤
	①	②	③	④	⑤
	①	②	③	④	⑤
	①	②	③	④	⑤
	①	②	③	④	⑤
	①	②	③	④	⑤
	①	②	③	④	⑤
	①	②	③	④	⑤
	①	②	③	④	⑤
	①	②	③	④	⑤
	①	②	③	④	⑤
	①	②	③	④	⑤
	①	②	③	④	⑤
	①	②	③	④	⑤
	①	②	③	④	⑤
	①	②	③	④	⑤
	①	②	③	④	⑤
	①	②	③	④	⑤
	①	②	③	④	⑤
	①	②	③	④	⑤
	①	②	③	④	⑤
	①	②	③	④	⑤
	①	②	③	④	⑤
	①	②	③	④	⑤
	①	②	③	④	⑤
	①	②	③	④	⑤
	①	②	③	④	⑤
	①	②	③	④	⑤
	①	②	③	④	⑤
	①	②	③	④	⑤
	①	②	③	④	⑤
	①	②	③	④	⑤
	①	②	③	④	⑤
	①	②	③	④	⑤
	①	②	③	④	⑤
	①	②	③	④	⑤
	①	②	③	④	⑤
	①	②	③	④	⑤
	①	②	③	④	⑤
	①	②	③	④	⑤
	①	②	③	④	⑤
	①	②	③	④	⑤
	①	②	③	④	⑤
	①	②	③	④	⑤
	①	②	③	④	⑤
	①	②	③	④	⑤
	①	②	③	④	⑤
	①	②	③	④	⑤
	①	②	③	④	⑤
	①	②	③	④	⑤
	①	②	③	④	⑤
	①	②	③	④	⑤
	①	②	③	④	⑤
	①	②	③	④	⑤
	①	②	③	④	⑤
	①	②	③	④	⑤
	①	②	③	④	⑤
	①	②	③	④	⑤
	①	②	③	④	⑤
	①	②	③	④	⑤
	①	②	③	④	⑤
	①	②	③	④	⑤
	①	②	③	④	⑤
	①	②	③	④	⑤
	①	②	③	④	⑤
	①	②	③	④	⑤
	①	②	③	④	⑤
	①	②	③	④	⑤
	①	②	③	④	⑤
	①	②	③	④	⑤
	①	②	③	④	⑤
	①	②	③	④	⑤
	①	②	③	④	⑤
	①	②	③	④	⑤

Mucosal Wave						
Left	①	②	③	④	⑤	
	Normal	Slightly Decreased	Moderately Decreased	Severely Decreased	Barely Perceptible	Absent
Right	①	②	③	④	⑤	

Non-vibrating Portion						
Left	①	②	③	④	⑤	
	None	20%	40%	60%	80%	100%
Right	①	②	③	④	⑤	

Phase Closure										
⑤	④	③	②	①	①	②	③	④	⑤	
Open phase predominates (whisper dysphonia)			Normal					Closed phase predominates (Glottal fry-extreme hyper-adduction)		

Phase Symmetry						
①	②	③	④	⑤		
Symmetrical	asymmetrical during begin or end tasks	asymmetrical during extremes of pitch or loud	asymmetrical during 50%+	generally asymmetrical	always asymmetrical	

Scale for rating phase symmetry on the SAV form. The scale is a mix of percentages of time and descriptive statements. The lack of a uniform scale may confuse raters.

Regularity						
①	②	③	④	⑤		
Regular	irregular during begin or end tasks	irregular during extremes of pitch or loud	irregular during 50%+	generally irregular 75%+	always irregular	

Scale for rating regularity on the SAV form. The scale is a mix of percentages of time and descriptive statements. The lack of a uniform scale may confuse raters.

B

FIG. 2 (continued) B. Diagram illustrating the design of the SAV form with comments about selected parameters.

Annexe 5 : Protocole du V.H.I.

Index des anomalies de la voix

Date :

Score : F : / 40

E : / 40

P : / 40

Total : / 120

Vous trouverez ci-dessous des phrases utilisées par la plupart des gens pour décrire leur voix et ses conséquences sur leur vie.

Entourez la réponse qui indique la fréquence à laquelle vous vivez cette expérience.

“Toujours” correspond à “oui, toujours un problème”;

“Jamais” correspond à “non, pas de problème”.

Merci de votre collaboration.

F1. On m’entend difficilement à cause de ma voix

Jamais Presque
 jamais Parfois Presque
 toujours
|-----|-----|-----|-----|

P2. Je suis à cours de souffle quand je parle.

Jamais Presque
 jamais Parfois Presque
 toujours
|-----|-----|-----|-----|

F3. On me comprend difficilement dans un milieu bruyant

Jamais Presque
 jamais Parfois Presque
 toujours
|-----|-----|-----|-----|

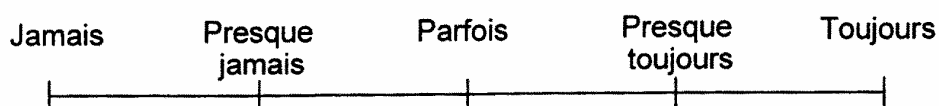
P4. Le son de ma voix varie en cours de journée

Jamais Presque
 jamais Parfois Presque
 toujours
|-----|-----|-----|-----|

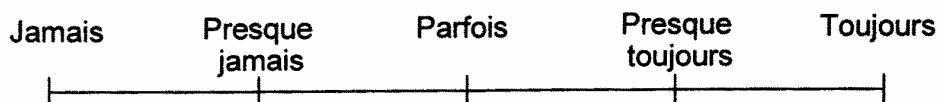
F5. Ma famille a du mal à m’entendre quand j’appelle dans la maison

Jamais Presque
 jamais Parfois Presque
 toujours
|-----|-----|-----|-----|

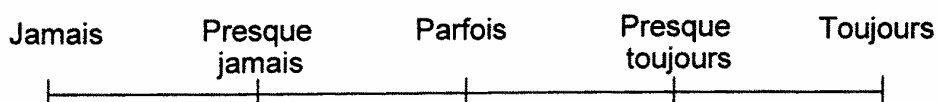
F6. Je téléphone moins souvent que je le voudrais



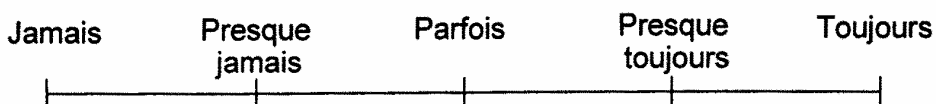
E7. Je suis tendu(e) quand je parle avec d'autres à cause de ma voix.



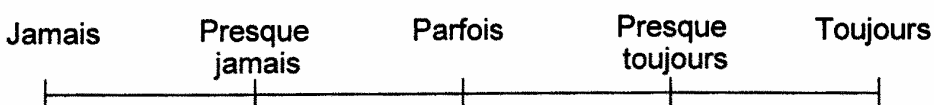
F8. J'essaie d'éviter les groupes de gens à cause de ma voix.



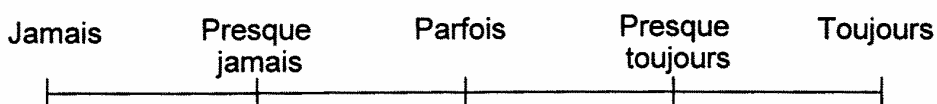
E9. Les gens semblent irrités par ma voix.



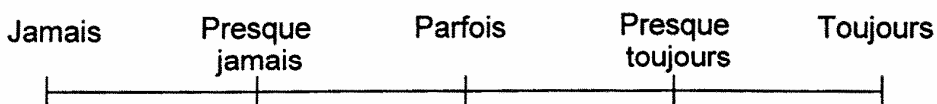
P10. On me demande : "Qu'est-ce qui ne va pas avec ta voix ?"



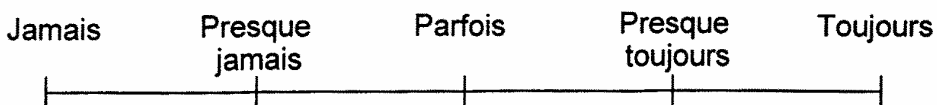
F11. Je parle moins souvent avec mes amis, mes voisins, ma famille à cause de ma voix.



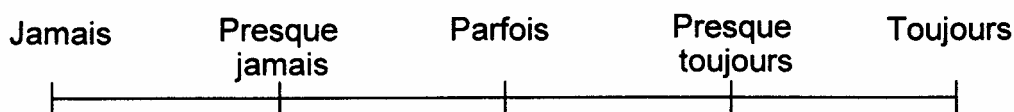
F12. On me demande de me répéter quand je dialogue face à face avec quelqu'un.



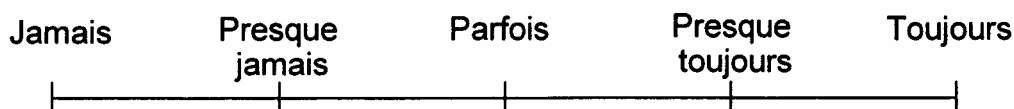
P13. Ma voix semble "cassante" et sèche.



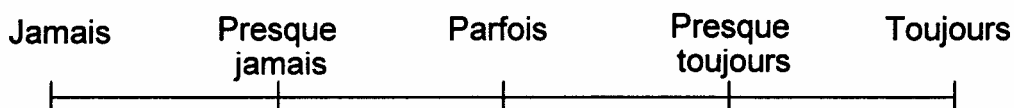
P14. J'ai l'impression que je dois forcer pour produire la voix.



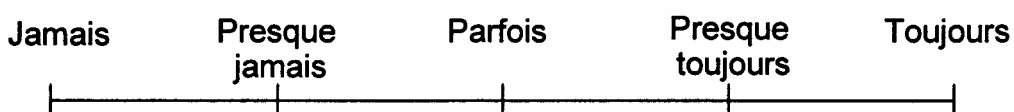
E15. Je trouve que les autres personnes ne comprennent pas mon problème de voix.



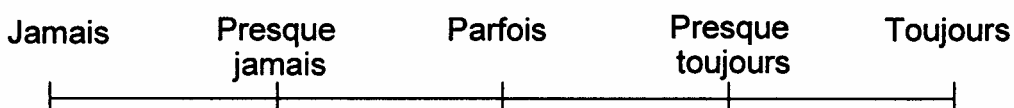
F16. Mes difficultés de voix limitent ma vie personnelle et sociale.



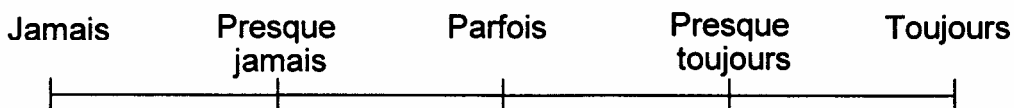
P17. La clarté de ma voix est imprévisible.



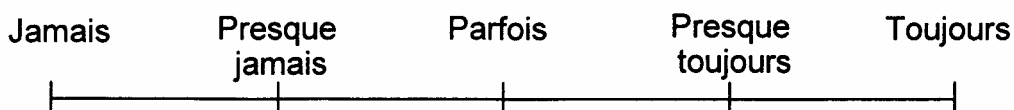
P18. J'essaie de changer ma voix pour qu'elle sonne différemment.



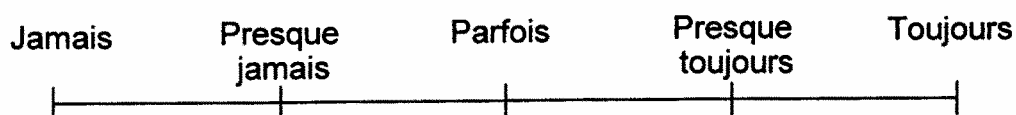
F19. Je me sens écarté(e) des conversations à cause de ma voix.



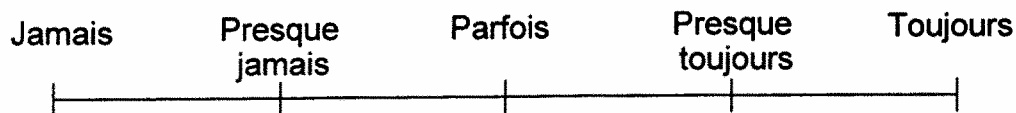
P20. J'ai l'habitude de faire beaucoup d'effort pour parler.



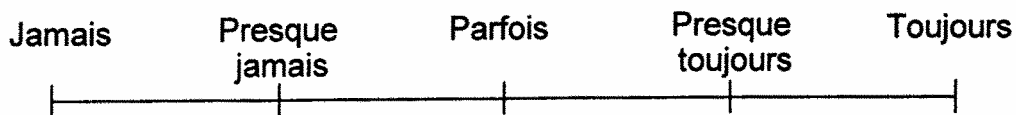
P21. Ma voix est plus mauvaise le soir.



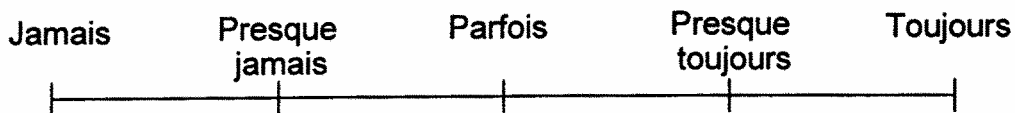
F22. Mes problèmes de voix entraînent des pertes de revenus...



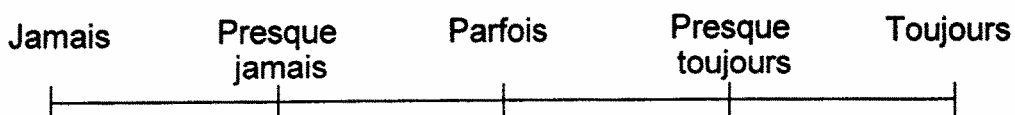
E23. Mon problème de voix me tracasse.



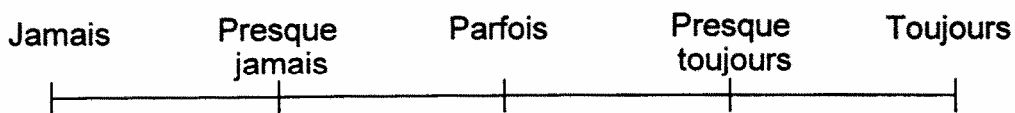
E24. Je sors moins à cause de mon problème de voix.



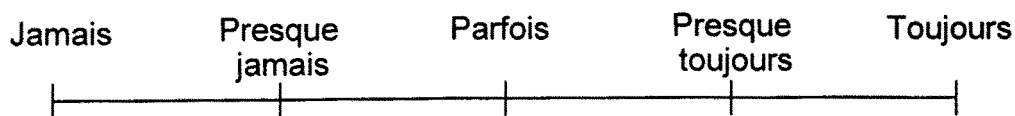
E25. Je me sens handicapé(e) à cause de ma voix.



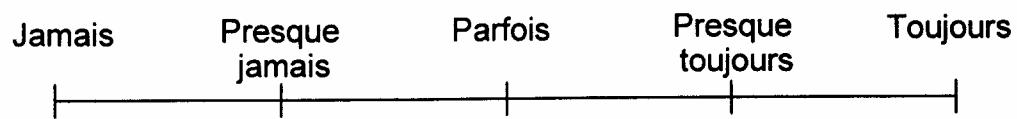
P26. Ma voix s'en va au milieu de la conversation.



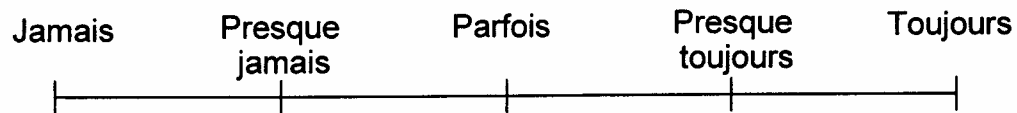
E27. Je suis ennuyé(e) quand les gens me demandent de me répéter



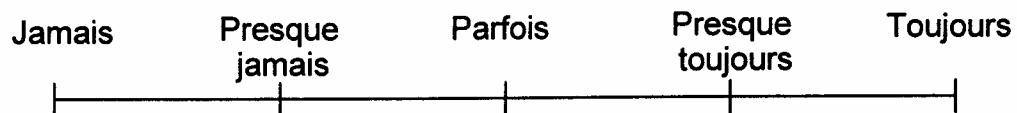
E28. Je suis embarrassé(e) quand les gens me demandent de me répéter.



E29. A cause de ma voix, je me sens incompetent.



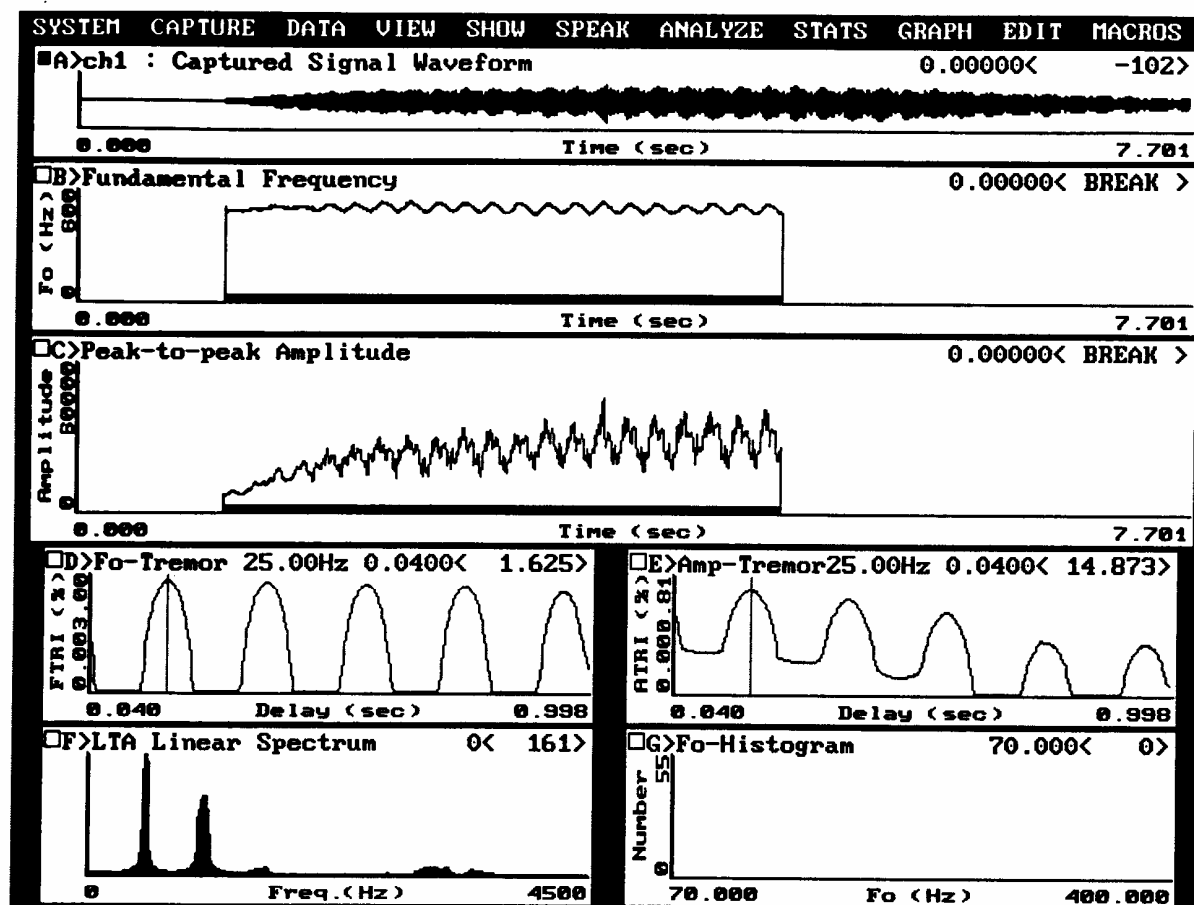
E30. Je suis honteux(se) de mon problème de voix.



Annexe 6 : Logiciels disponibles sur le marché.

1. **PCLX**, Laryngograph Ltd. 1 Foundry Mews, London NW1 2PE, UK.
2. **Phonedit**, SQ lab., C/CEEI, Domaine du petit Arbois – BP88, 13545 Aix-en-Provence cedex a, France. <http://www.sqlab.com/>
3. **Phonos 1** (communiqué par Bourgogne – Nguyen Vantrieu Ch, 1999 (Université de Lille II)).
4. **Signalyse**, Software Manager, Phonetics lab., Linguistics Dept, UCLA, LoCA 90024-1543, Prof. Keller E, Laboratoire d'analyse informatique de la parole (LAIP), Lettres, Université de Lausanne, CH-1015 Lausanne, Switzerland.
5. **Spectra plus**. <http://www.soundtechnology.com/home.htm>
6. **N-Unice** (communiqué par Bourgogne – Nguyen Vantrieu Ch, 1999 (Université de Lille II))
7. **O-Winsnorri** version 1.02, Laboratoire Lorrain de recherche en informatique. Université Henri Poincaré- Nancy I, BP 239, 54506 – Vandoeuvre les Nancy Cedex. <http://granvelle.univ-fcote.fr/phonetic/techniques/logiciels.htm>.
8. **C.E.C.I.L.** (Computerised Extraction of Components of Intonation in Language). <http://www.jaars.org/icts/software.htm>.
9. **Dr Speech**, <http://www.drspeech.com/information.html>
10. **Phoneval** , API, 31 rue Bellecombe 69006 Lyon.
11. **Soundscope**, GW Instruments, 35 Medford street, somerville, MA 02143, USA.

Annexe 7 : Protocole du M.D.V.P.



Code: 15 a.

Année Naissance: 1964

Catégorie: pro.

Tessiture: soprano lyrique léger

Son(s) produit(s): a vibré.

Sampled Data: not named
 Signal Level: 27080
 Sampling Rate: 50000
 Time Range: 0.00000 sec
 7.70136 sec

Average Fundamental Frequency	Fo	=	518.403	Hz
Average Pitch Period	To	=	1.931	ms
Highest Fundamental Frequency	Fhi	=	554.631	Hz
Lowest Fundamental Frequency	Flo	=	486.618	Hz
Standard Deviation of Fo	STD	=	15.936	Hz
Phonatory Fo-Range in semi-tones	PFR	=	4	
Fo-Tremor Frequency	Fftr	=	5.263	Hz
Amplitude Tremor Frequency	Fatr	=	5.263	Hz
Length of Analyzed Sample	Tsam	=	7.701	s
Absolute Jitter	Jita	=	8.303	us
Jitter Percent	Jitt	=	0.430	%
Relative Average Perturbation	RAP	=	0.228	%
Pitch Perturbation Quotient	PPQ	=	0.256	%
Smoothed Pitch Perturb. Quotient	sPPQ	=	1.131	%
Fundamental Frequency Variation	vFo	=	3.074	%
Shimmer in dB	ShdB	=	0.289	dB
Shimmer Percent	Shim	=	3.392	%
Amplitude Perturbation Quotient	APQ	=	4.421	%
Smoothed Ampl. Perturb. Quotient	sAPQ	=	9.498	%
Peak-Amplitude Variation	vAm	=	30.793	%
Noise to Harmonic Ratio	NHR	=	0.0856	
Voice Turbulence Index	VTI	=	0.0415	
Soft Phonation Index	SPI	=	10.2974	
Fo-Tremor Intensity Index	FTRI	=	2.879	%
Amplitude Tremor Intensity Index	ATRI	=	18.592	%
Degree of Voice Breaks	DVB	=	0.000	%
Degree of Sub-harmonics	DSH	=	0.000	%
Degree of Voiceless	DUV	=	13.115	%
Number of Voice Breaks	NVB	=	0	
Number of Sub-harmonic Segments	NSH	=	0	
Number of Unvoiced Segments	NUV	=	32	
Number of Segments Computed	SEG	=	244	
Total Pitch Periods Detected	PER	=	1999	

Code: 15 a.
 Année Naissance: 1964
 Catégorie: pro.
 Tessiture: soprano lyrique léger
 Son(s) produit(s): a vibré.

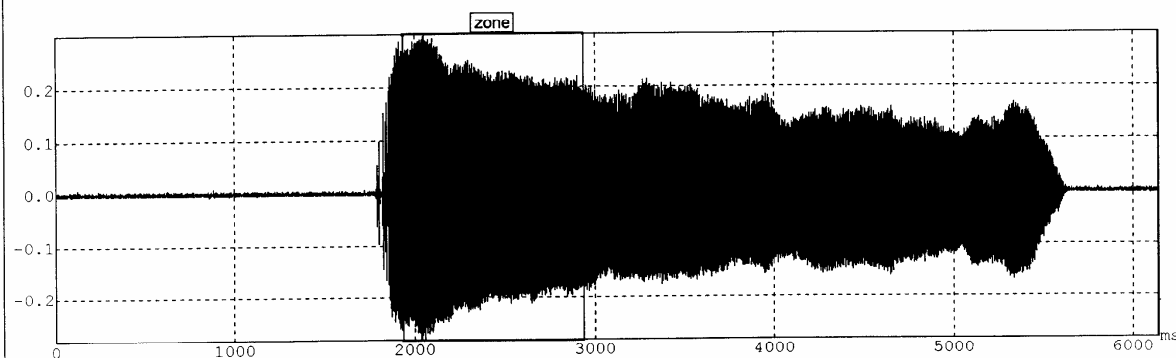
Annexe 8 : Protocole de E.V.A.

Voice Profile

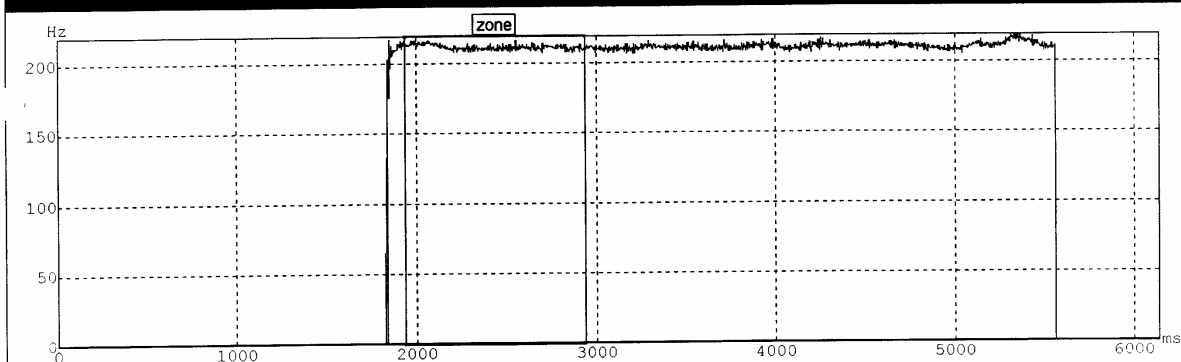
Record page 1/3

[16-7-1999] Charlotte : Normal

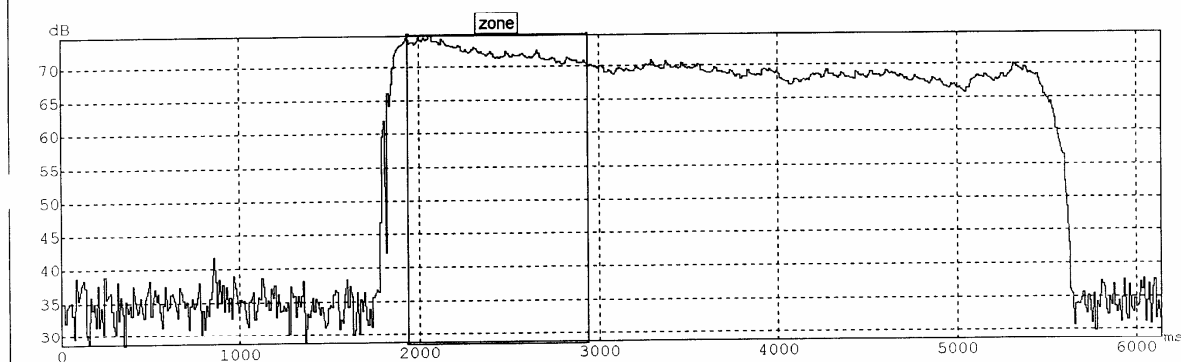
Speech



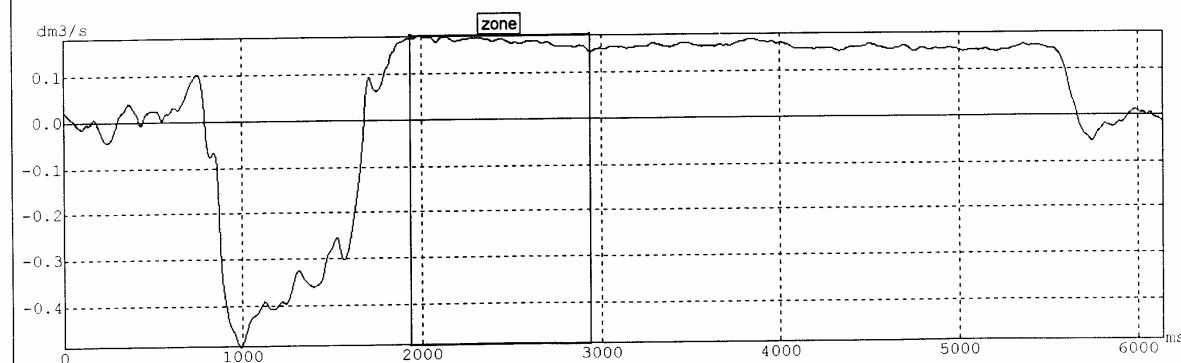
Fundamental frequency



Intensity



Oral airflow



O.R.L. - LOGOPEDIE INTERNE

Voice Profile

Record page 2/3

[16-7-1999] Charlotte : Normal

	fundamental frequency	intensity
mean	212.01 Hz	72.16 dB
std dev	2.05 Hz	1.16 dB
c.v	0.97 %	1.60 %
CTC perturbation	2.056 Hz (jitter)	0.264 dB (shimmer)
average perturbation	0.0062 (RAP)	0.03716 (APQ)
percentage	0.97 % (jitter factor)	0.34 % (Shimmer factor)
other	9.71 (jitter ratio)	

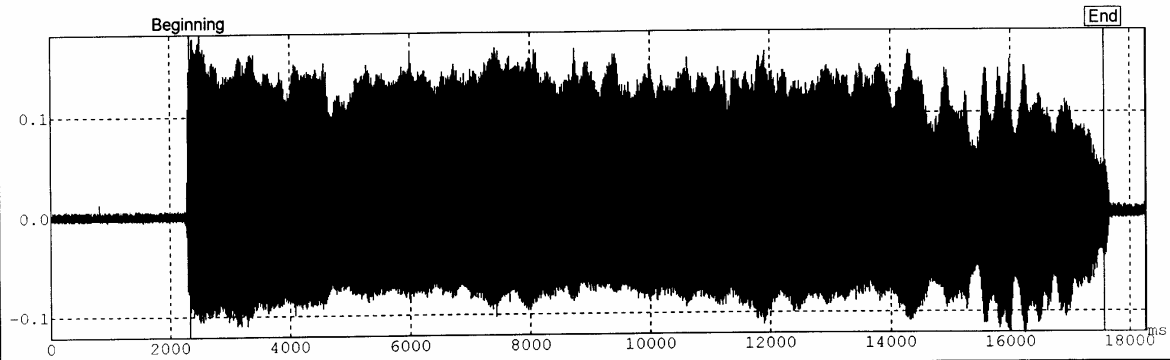
	F0	sr	NNE	HNR	sr hf	NNE hf	HNR hf
[1936.5,2936.5]	211.3	92.1	-11.04	11.70	78.0	-6.58	3.55

	glottic leakage	mean	std dev	c.v
Areodynamic results	2.34 cc/s/db	0.169 dm3/s	0.008 dm3/s	4.58 %

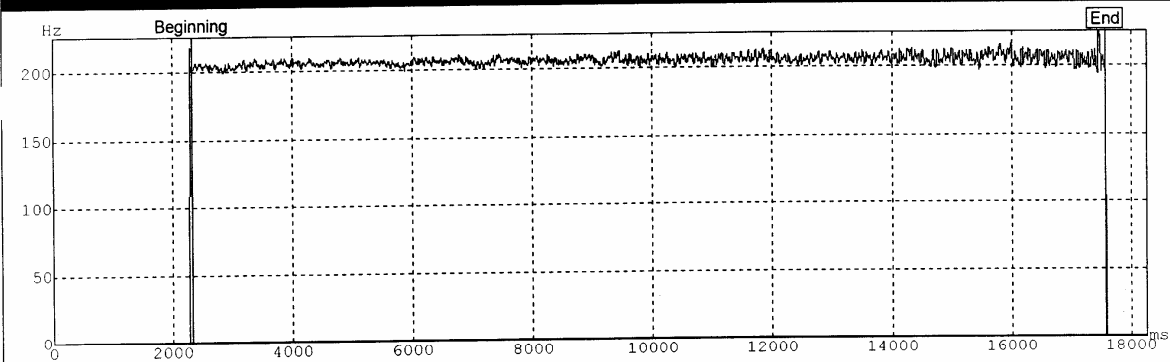
O.R.L. - LOGOPEDIE INTERNE

[16-7-1999] Charlotte : Normal

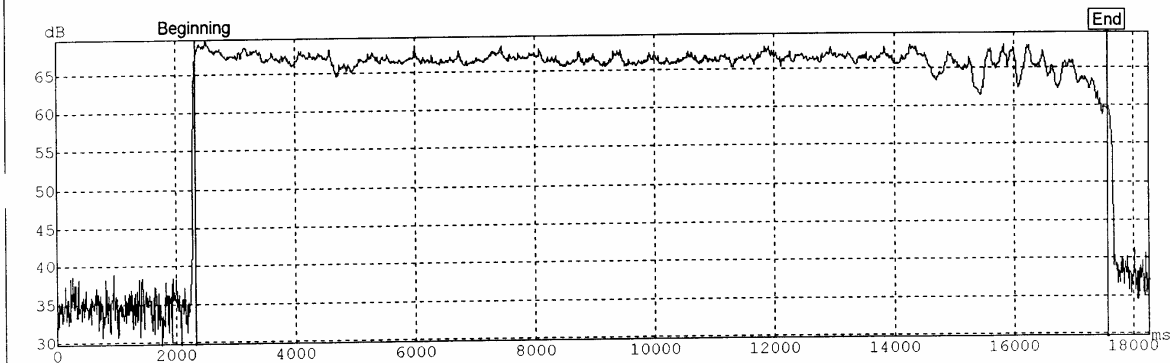
Time Length =15.245 sec.



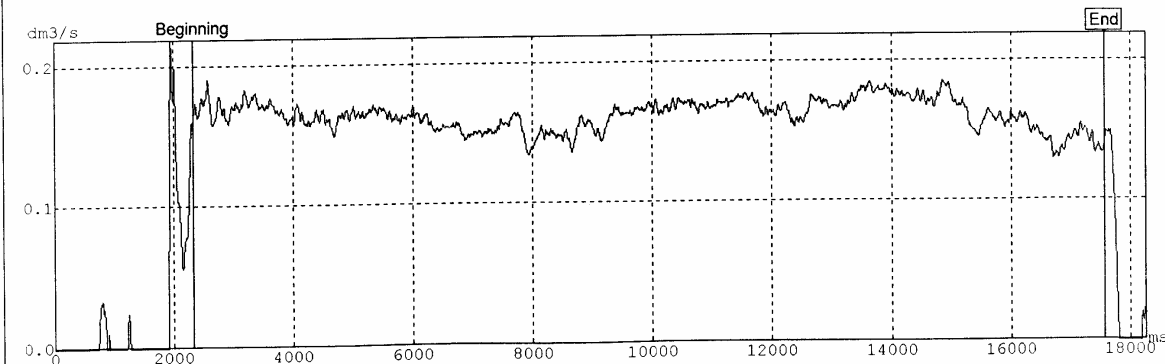
Mean F0=206.3 Hz



Mean Intensity=66.3 dB



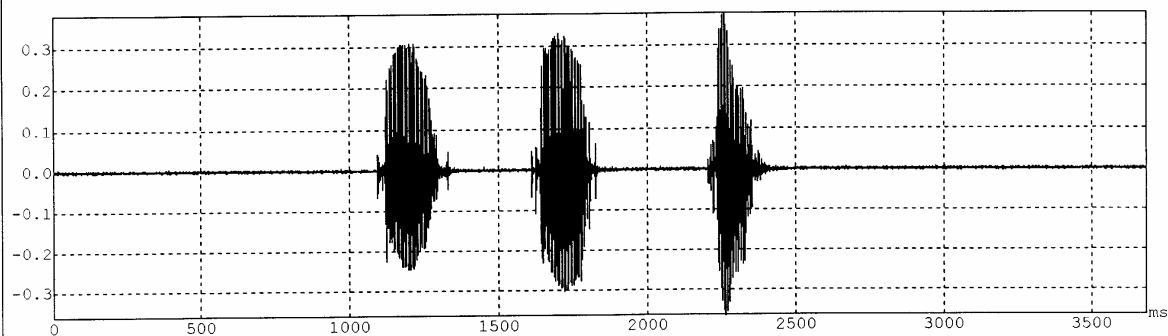
Exhaled Volume=2.475 dm3 Mean Airflow = 0.162 dm3/s



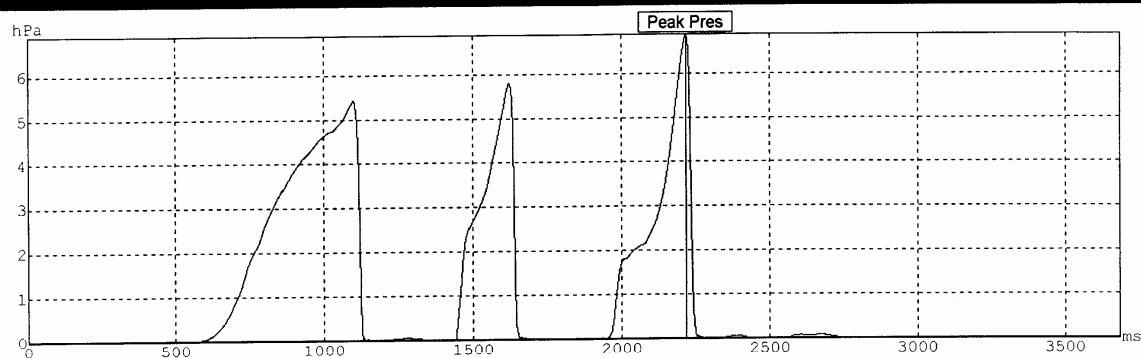
O.R.L. - LOGOPEDIE INTERNE

[16-7-1999] Charlotte : Normal

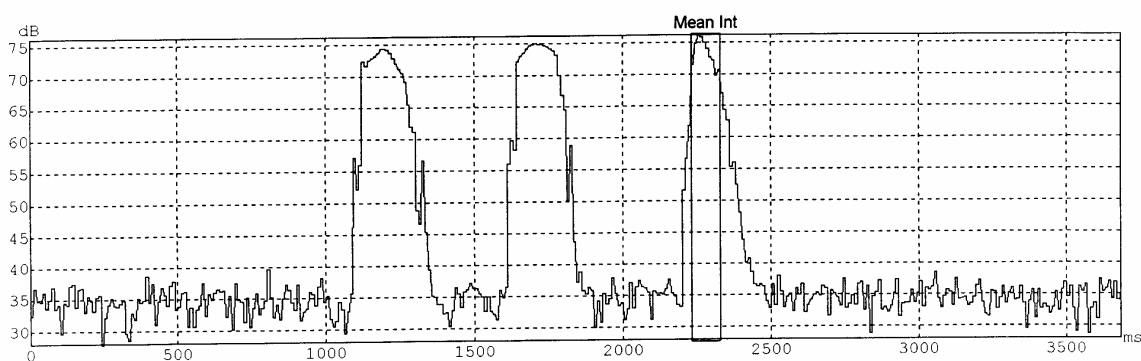
Speech



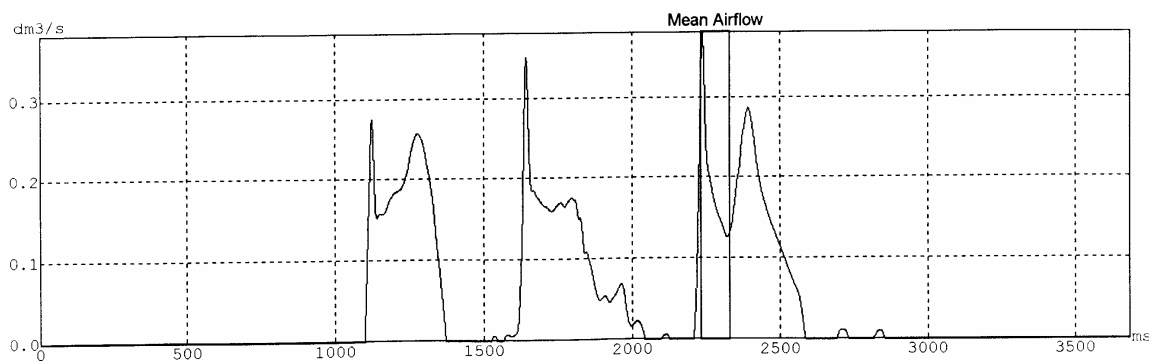
Pressure t=2216.6 v=6.9 hPa



Mean Intensity t=[2229.8;2330.1] v=73.3 dB



Mean Airflow t=[2229.8;2330.1] v=0.19 dm3/s



O.R.L. - LOGOPEDIE INTERNE

Annexe 9 : Instructions données aux juges avant l'évaluation des 240 échantillons vocaux chantés.

Vous entendrez 240 échantillons musicaux : ils proviennent de chanteurs tous niveaux confondus. Leur tâche se répartit comme suit :

- Un [a] tenu vibré,
- Un [a] tenu non vibré,
- Un morceau au choix,
- Un morceau imposé.

Vous recevrez une grille d'évaluation par échantillon. Nous vous demandons de juger le vibrato et uniquement le vibrato, de ces échantillons. Nous insistons sur le fait qu'il ne s'agit pas d'émettre un jugement par rapport à la voix, à l'interprétation ou à un autre facteur de la qualité vocale.

Pour éviter toute saturation auditive, nous vous proposons d'étaler les auditions en 6 séances ou plus si nécessaire.

Avant d'écouter les enregistrements, il nous semble important de nous accorder sur la conception et la terminologie du vibrato, du tremolo, du chevrottement et du son blanc.

- Le **vibrato** consiste en une fluctuation périodique de la *fréquence fondamentale*, de l'*intensité* et du *timbre* du son (Castellengo (1991)).

Il donne à la voix une richesse expressive, une tendresse et une souplesse.

La périodicité est un phénomène qui se reproduit de manière identique dans le temps.

La fréquence fondamentale est la hauteur tonale du son. Elle s'exprime en hertz (Hz). Elle est déterminée par le nombre de vibrations par seconde ($la_3 = 440$ Hz). Si nous nous exprimons en qualité perceptive, nous disons du son qu'il est grave ou aigu.

L'intensité dépend de l'amplitude du son. En qualité perceptive, le son est décrit comme fort ou faible.

Le timbre est fonction du nombre et de l'intensité des harmoniques (composante spectrale), et de la transformation du son dans les cavités de résonance. En qualité perceptive, nous parlerons de clair ou sombre.

- Le **tremolo** ne possède pas la régularité du vibrato. Sa pulsation est trop rapide et irrégulière en amplitude. Sa perception est désagréable, gênante.
- Le **chevrotement** est un changement de hauteur tonale avec périodicité et amplitude larges et inégales.
- Le **son blanc** est un son détimbré (sans vibrato).

TABLEAU RECAPITULATIF.

	Paramètres acoustiques	Qualités perceptives
Hauteur	Fréquence	Grave/aigu
Intensité	Amplitude	Fort/faible
Timbre	Composante spectrale	Sombre/clair